Wuhong Wang
Klaus Bengler
Xiaobei Jiang *Editors*

Green Intelligent Transportation Systems

Proceedings of the 7th International Conference on Green Intelligent Transportation System and Safety

Springer

Lecture Notes in Electrical Engineering

Volume 419

About this Series

"Lecture Notes in Electrical Engineering (LNEE)" is a book series which reports the latest research and developments in Electrical Engineering, namely:

- Communication, Networks, and Information Theory
- Computer Engineering
- Signal, Image, Speech and Information Processing
- Circuits and Systems
- Bioengineering

LNEE publishes authored monographs and contributed volumes which present cutting edge research information as well as new perspectives on classical fields, while maintaining Springer's high standards of academic excellence. Also considered for publication are lecture materials, proceedings, and other related materials of exceptionally high quality and interest. The subject matter should be original and timely, reporting the latest research and developments in all areas of electrical engineering.

The audience for the books in LNEE consists of advanced level students, researchers, and industry professionals working at the forefront of their fields. Much like Springer's other Lecture Notes series, LNEE will be distributed through Springer's print and electronic publishing channels.

More information about this series at http://www.springer.com/series/7818

Wuhong Wang · Klaus Bengler
Xiaobei Jiang
Editors

Green Intelligent Transportation Systems

Proceedings of the 7th International Conference on Green Intelligent Transportation System and Safety

Springer

Editors
Wuhong Wang
Beijing Institute of Technology
Beijing
China

Klaus Bengler
Technical University of Munich
Munich
Germany

Xiaobei Jiang
Beijing Institute of Technology
Beijing
China

and

Technical University of Munich
Munich
Germany

Contents

A New Method for 3D-Simulating Traffic Accident Based on VISSIM

Shi-bo Zhang, Lan Liu, Hui Li and Ping-fei Li

Abstract A new method (VVSC method) for 3D-simulation of traffic accident was proposed, which was based on VISSIM platform frequently-used in traffic flow simulation. The general technical route was introduced followed by discussing modeling method for five key issues including: 3D element modeling, system framework construction, vehicle trajectory and process controlling, crash effect design and environment design. In the end, a successful case with complicated conditions was presented to show the application of the new method. The key point of VVSC method was making full use of both VISSIM and its secondary development interface COM. VVSC method is suitable for forensic institutions to simulate traffic accidents with accurate, fast, and cheap style.

Keywords Traffic engineering · Traffic accident · 3D simulation · VISSIM · Accident reconstruction

Introduction

In the process of handling and analysis of traffic accidents, 3D simulation animations are always expected to demonstrate accident by the concerned parties such as police, media, involved parties, and the public, because 3D simulation could improve the convincingness and representability of accident reconstruction results. At present, some 3D software such as 3DMAXP [1] can create such animation, but this kind of methods need professionals, more time and more cost, this is not fit for the field of accident identification which need low cost and high efficiency;

S. Zhang (✉) · L. Liu
College of Transportation and Logistics, Southwest Jiaotong University,
Chengdu 610031, China
e-mail: zhshbox@163.com

S. Zhang · H. Li · P. Li
School of Automobile and Transportation, Xihua University, Chengdu 610039, China

moreover, some other software such as PC-Crash [2], Madymo can reconstruct accident with simple animation which is lack of representability.

The new method, proposed in this paper, will make use of COM interface by means of secondary programming, and based on VISSIM platform which is frequently used in traffic flow simulation. The aim of the new method is to make sure that accident engineers make full use of their knowledge of accident reconstruction, focus on the analysis of accident process, thereby, complete accident simulation fast, accurately, and modularly. This hasn't been reported in the current literatures so far.

General Technical Route

VISSIM [3], which is based on time interval and driving behaviors, is a microscopic traffic flow simulation system developed by PTV Company. The regular function of VISSIM is to simulate and analyze traffic flow under various traffic conditions [4]. For accident simulation, which needs describing the individual crash movements and effects, VISSIM has no special considerations when it was designed. However, the authors of this paper "discover" that VISSIM could be used to simulate accident, because VISSIM has a secondary programming interface which allows users to control the underlying information of simulated objects, and then achieve the customized effects and movements.

In 3D-simulation of traffic accidents, many elements should be included such as vehicles, pedestrians, road facilities, trees, terrain, etc. These elements can be controlled by programming in COM interface. However, VISSIM can't produce a new element by itself; so, some other tools are used to model static objects. The format of 3D models compatible for VISSIM includes .v3d and .skp, which are created by SketchUP [5] and V3DM [6] respectively. SketchUP is free 3D design software developed by Google Company. The user could get a great deal of source materials from open library, which can be improved easily and imported into VISSIM as static elements in accident environment. V3DM is a simple 3D tool for modeling vehicles and buildings developed by PTV Company.

Figure 1 is the general technical route based on VISSIM, V3DM, SketchUP and COM interface. The new method could be named "VVSC method" from the first letters of VISSIM, V3DM, SketchUP and COM.

Modeling Method for Key Issues

Modeling 3D Element

There are three types of 3D element in traffic accident simulation as follows:

Fig. 1 General technical
route of VVSC method

Simple 3D models based on V3DM	Complicated 3D models based on Sketch Up	Traffic flow modeling based on VISSIM

3D state definition based on V3DM

VB programming to design the special movements of single 3D object based on COM interface

Coordinate the movements of all 3D objects; Add pedestrians, bicycles, road facilities, buildings, etc.

Run simulation, export and edit video

Type-A element: ① Road pavement and markings, such as lane dividing lines, crosswalk lines and oriented arrows of vehicle lanes, non-vehicle lanes and pedestrian lanes. ② Traffic facilities such as traffic signals, street lamps, traffic signs. ③ Road constructions such as bridges, tunnels, interchanges, toll gates, bus stops. ④ Buildings and road plantings such as trees, lawn.

Type-B element: the concerned individuals involved in the accident with specific shape and logo.

Type-C element: the background traffic flow including vehicles, pedestrians, bicycles, etc., which aren't involved in the accident.

The built-in models in VISSIM model library are .v3d format, including common automobiles, planes, pedestrians, etc., which have foreign surface paintings, and the users can't modify them. So, the built-in models are just fit for Type-C element. For Type-A and Type-B elements, SketchUP and V3DM could be used to construct by means of three methods as follows:

(1) Completely creating based on V3DM

The project file of V3DM is .vvp format which can be modified, this means the user can construct any shapes, any paintings, and any logos, etc. This method is fit for objects with simple and angular shapes such as road central divider and conical bucket (Fig. 2). Most of Type-A element are recommended to be constructed by means of this method. More details for modeling in V3DM could related to the User's Guide.

(2) Direct use (or little modification) of SketchUP models from open library

In order to make lifelike simulation, complicated 3D models, i.e. trees, bridges, etc., are also needed besides simple models. It is very complex (even impossible) to construct these models in V3DM. Actually, most complicated models can be obtained from open library of SketchUP with .skp format, and put onto certain

Fig. 2 Example model of
V3DM
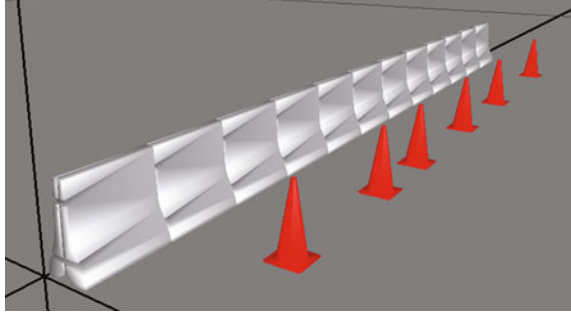
Fig. 3 Example model of
SketchUP
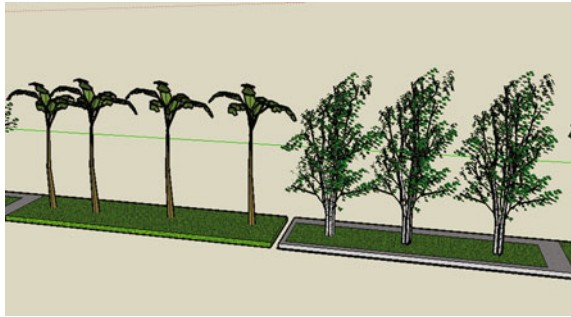
position in simulation environment. The library has massive models for free, users
can certainly find out useful models in the library. If there is no identical model
needed by users, similar model can also be chosen to modify simply (Fig. 3).

(3) SketchUP model plus V3DM splice and painting

For some elements with specific shape and painting (it is important for accident
simulation), the most effective method is to splice small components of SketchUP
model together and paint in V3DM. Figure 4 is an example of SketchUP plus
V3DM.

Framework for Simulation System

The crucial reason why VISSIM can be used to simulate traffic accident is COM
interface, which is rarely used in the common application. VISSIM can be applied
into other application programming interface via COM, the models and underlying
messages (object, method, attribute) can also be called interactively via COM. User
can apply all RAD (Rapid Application Development) tools, including Visual Basic
Script, Java Script or Visual C++, Visual J++, etc.
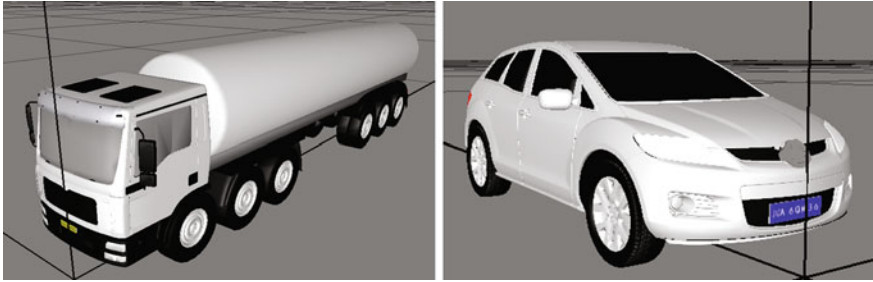
Fig. 4 Example models of SketchUP model plus V3DM splice and painting

Fig. 5 Programming interface of VISSIM-COM
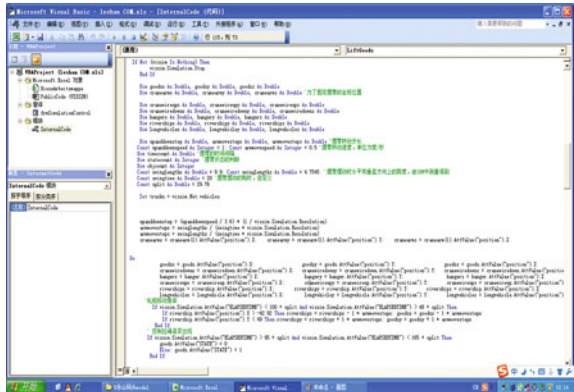
Fig. 6 User interface of simulation system
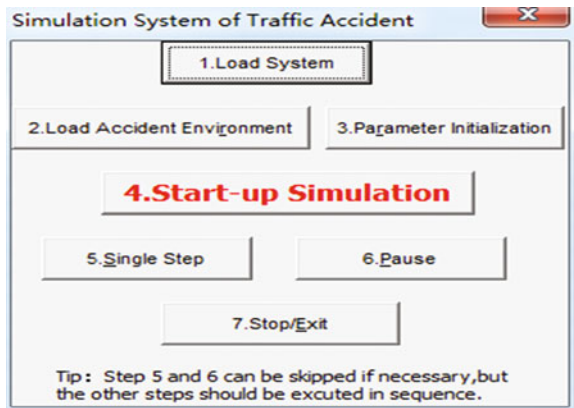
Colloquially, VISSIM-COM provides a "vissim" library, which can be called and controlled by VB or VC etc., there are 14 kinds of "vissim" object including link, node, parkinglot, path, vehicle, staticobject etc., which has strict ranks.

Fig. 7 Photo I of case accident scene

Fig. 8 Photo II of case accident scene

Fig. 9 Screenshot I of case accident simulation

"vissim" is the top object with the highest rank, other objects are the subobject of "vissim". Visual Basic has special language "For Each … Next" in order to call "vissim" objects in proper sequence. Additionally, every object has its own methods and attributes, i.e. Object "Vehicle" has "SPEED", "POINT", "LANE" and other 34 attributes, "AddVehicleAtLinkCoordinate" (add a new car on the current link), "RemoveVehicle" and other 9 methods.

Visual Basic was selected to be programming interface in the traffic accident simulation system developed by the authors (Fig. 5). Figure 6 is the user interface of the system based on the general framework with many controlling languages to arrange simulation process as follows:

---

① Load traffic accident simulation system:

Set vissim = CreateObject("VISSIM.Vissim")
Set guiSheet = Worksheets("VISSIM")

② Load accident environment:

path = Application.ThisWorkbook.path
filename = guiSheet.Range("Filename")
vissim.LoadNet (path + filename), 0
guiSheet.Project_InitMain

③ System initialization:

workdir = Application. ThisWorkbook.path
longvehicle.AttValue("position") = vissim.NewWorldPoint(−67,   37.8, −10)

④ Continuous simulation:

Do

RunOneTimeStep
DoEvents

Loop   While   Not   singleStep   And   (vissim.Simulation.AttValue ("ISRUNNING"))

⑤ Single-step simulation: *RunOneTimeStep*
⑥ Simulation Pause: *vissim.Simulation.Stop*
⑦ Simulation Stop/Exit: *vissim.Exit*

Based on the framework above, put the languages of controlling the movement of concerned objects into the "Start-up" module, the simulation will be true.

Vehicle Trajectory and Process Control

There are two ways to present trajectory in VISSIM: ① Draw "link" route, and make the customized vehicle appear on the route, and the vehicle will travel along the route until the end of the route (Method-A); ② Real-time change the coordinates of vehicle objects by means of COM programming (Method-B). It is easy to operate via Method-A, and the attributes (e.g. length) of "link" could easily be used as determinant of vehicle state. In the meanwhile, it is complex to design algorithms of coordinate transformation via Method-B. So, Method-A is recommended.

In addition, the initial time and position of vehicles should be programmed via COM method "AddVehicleAtLinkCoordinate". For example, "*Set truck = trucks. AddVehicleAtLinkCoordinate(700,7,33,1,92, 0)* "means that a "truck" of vehicle 700 (vehicle type code) will be added at link 33, lane 1, 92 m length, with speed of 7 km/h. Vehicle type code should be preassigned in VISSIM. To erase a certain vehicle in simulation environment, "RemoveVehicle" could be used. For example, "*trucks.RemoveVehicle (20)*" means a "object" of code 20 in "trucks" set will be removed from VISSIM.

After planning vehicle trajectory, it is needed to adjust emerging time and speed rhythm to coordinate each other so as to like the real accident. So, some information, such as crash type, crash position, crash speed, and acceleration, etc., should be obtained in advance. In COM interface, the program to control the process of vehicles is similar with the following languages:

——

If veh1.AttValue("totaldistance") > 620 Then veh1.AttValue("speed") = veh1. AttValue("speed")−0.1

——

It means when the travel distance of vehicle "veh1" is larger than 620, the speed will reduce 0.1 every time step.

Figures 10 and 11 shows a case simulation with vehicle trajectory and process control.

Crash Effects Design

Accident crash is always a process of energy conversion, which is always sudden change finished in less than one second. The change may result in scattered objects, deformation, rotation, rollover, disintegration, fire, or dust diffusing, etc. The means to achieve these effects in VISSIM is usually related to the functions of state control and coordinate control for static object via COM. Unlike the vehicles traveled on

Fig. 10 Screenshot II of case accident simulation

Fig. 11 Screenshot III of case accident simulation

the "link", static objects in simulation environment are placed as "StaticObject" manually. Every movement and state change of static object should be programmed in COM interface; of course, if the object would keep rest, no program would needed. The methods to control static object are as follows:

───

① Load static 3D object. e.g.
Set guanche = StaticObjects.GetStaticObjectByName(path + "\ guanche. skp")
means: set "guanche.skp" model in the path of "path" to "guanche" as static object.

Fig. 12 Screenshot IV of case accident simulation
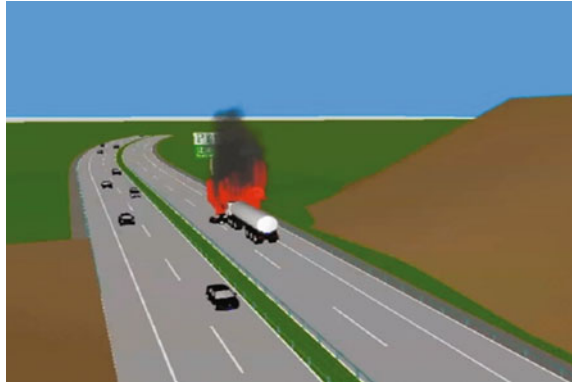
② Change the state of 3D object. e.g.
  guanche.AttValue("3DMODELSTATE") = 1
  means: set the state of object "guanche" to No.1 state.
  the states of 3D object should be set up in V3DM, a object can have a variety of state.
③ Change the coordinate of 3D object. e.g.
  house.AttValue("position") = vissim.NewWorldPoint(0, 0, −8)
  means: set the coordinate of object "house" to (0, 0, −8).

_____

The change of coordinate can control emerging or not, moving or not, fast or slow, etc. Only deal with state control and coordinate control, the crash effects would be good in accident simulation. Figure 12 shows a case simulation with fire effect.

Traffic Environment Design

Traffic environment for traffic accident includes accident scene and surrounding except the directly concerned parties, such as road, facilities, trees, buildings and pedestrians, bicycles, traffic flow, etc., which determine the realistic degree of 3D effect. For static environment, the method introduced above could be used to construct static objects, which should be exported into simulation environment, followed by move, rotating, zoom, etc. For the dynamic environment, the conventional function of traffic simulation of VISSIM could be used, more details are related to the literatures. Figure 9 shows a case simulation with traffic environment.

Case Application

December 8 2013 15 pm, a semi-trailer towing vehicle, which traveled on Chengzilu expressway from Zigong to Chengdu at K95 + 500 m, crashed with 16 other vehicles, resulted in 7 persons died, 23 persons injured (Figs. 7, 8). Because of the complexity and social impact of the case, the police delegated the authors to reconstruct the accident. In the meantime, because there were too many reasons (too many vehicles, too long distance, fire, scattered objects, deformation, rotation, rollover, disintegration, dust diffusing) not easy to deal with, several regular methods (such as PC-CRASH, MADYMO, ARAS, 3DMAX) were incapable to do.

The new method (VVSC method) was applied to reconstruct the accident, and it was successful to simulate the whole process including the most of crash effects, and form a film with various angles. The simulation helped the police to handle the case, and increase the understanding for the concerned parties and public. Figures 9, 10, 11, and 12 are the screenshots of the simulation.

Conclusions

3D-simulation technologies have been needed and used more and more in the analysis of traffic accident, and various methods are also needed to deal with various accident situations. As a new method for accident simulation, VVSC method has several advantages as follows: ① The method can make full use of the open free 3D model library, not only shorten production cycles, but also reduce the cost; ② The method is flexible to fix crash time, crash position, deformation, and state of 3D objects because of excavating the underlying information of VISSIM via COM interface; ③ The method is fit for traffic engineers, not fit for 3D engineers, because the method was developed based on the background of traffic engineering. Therefore, the method is suitable for forensic institutions to simulate traffic accidents with accurate, fast, and cheap style.

References

1. Peng Guobin, He Yueqing, Sun Yu, and Zhou, Kai Xi. 2011. Three-dimensional game modeling and design research based on 3dmax software. *Communications in Computer and Information Science* 215(2): 192–196.
2. Zhang Yong Gang, Xu Jian Min, Zou Tie Fang, and Liu Y. 2014. A method for reconstructing vehicle-vehicle impact accidents based on Pc-Crash. *Applied Mechanics and Materials* 641–642: 799–804.

3. PTV VISSIM. 2014. User Manual. Germany: PTV Group.
4. Shi Junqing, and Cheng Lin. 2013. Simulation and analysis of highway traffic accident based on VISSIM. *Applied Mechanics and Materials* 253–255(1):1682–1685.
5. Ma Liang, Wang Fen, and Bian Hai. 2012. SketchUp Pro 8.0 Tutorial. Beijing: Posts & Telecom Press.
6. V3DM. 2013. User Manual. Germany: PTV Group.

True 3D Surface Feature Visualization Design and Realization with MapGIS K9

Yuan Yuan, Shao Chun-fu, Ji Xun, Xiang Huaikun and Zhang Wenji

Abstract In the True 3D GIS system, obtaining and modeling of 3D geometric data is the key problem. This paper studies the concept, steps, and method of designing and realizing true 3D visualized campus map with MapGIS K9 software, Daogle Google interceptor, 3DMAX, Photoshop, enhancing the acquisition speed of 3D geometric data and solving the problems such as integration of MapGIS K9 software and 3DMAX software models with Shenzhen Polytechnic as an example to realize the 3D visualized campus.

Keywords True 3D GIS · MapGIS K9 · 3DMAX · Daogle Google interceptor · Map

Introduction

True 3D Volumetric Display Technique is a computer stereo display technology, which can directly observe the 3D images with physical depth of field. Since the true 3D model is nearly identical with the reality environment, it can bring a strong and life-like sensory shock to the users so that they can feel personally on the scene. Users can roam freely in a true 3D scene. The true 3D GIS takes the 3D space

Y. Yuan · S. Chun-fu (✉) · J. Xun
MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology,
Beijing Jiaotong University, Beijing 100044, China
e-mail: cfshao@bjtu.edu.cn

Y. Yuan
Department of Civil and Structural Engineering, The Hong Kong Polytechnic University
Hung Hom, Kowloon, Hong Kong

X. Huaikun
School of Automotive and Transportation Engineering, Shenzhen Polytechnic (East
Campus), Xili, Nanshan District, Shenzhen City 518055, China

Z. Wenji
Zondy Cyber, Wuhan City 430073, China

coordinates (x,y,z) as independent parameters to conduct geometric modeling for spatial entity objects, whose mathematic expression is $F = f\,(x,y,z)$. The model established in this way not only can realize true 3D visualization, but also can be used for 3D space analysis [1]. It indicates that the true 3D GIS is more realistic than previous 2D and 2.5D, which can observe the objects from each direction. Meanwhile, the true 3D GIS is also equipped with the functions of common GIS like collecting, storing, managing, analyzing, and displaying the spatial and nonspatial data.

Currently, there are several trends in the production of 3D scenes. In the first trend, ARCGIS and SketchUP are used for modeling based on WGS1984 coordinate system. Total station +RTK is used to measure and acquire the surface feature plan in the format of CASS CAD and the data of surface elevation points. Meanwhile, remote elevation measurement is employed to acquire the data of top elevation of surface features [2]. Though the plan data of surface features can be acquired from the digitized mapping with a total station, the method is time consuming and heavy in work load. Besides, ARCGIS is famous for long in the industry as a foreign software, but it is much more expensive than MapGIS K9, a domestic software with similar functions. In the second trend, the vehicle-mounted laser scanning system and GPS are adopted at the surface to figure out the outline and geographic data of surface features through ranging. The advantage is that the speed to acquire information is fast, the geometric information acquired is rather precise. The disadvantage is the work load is heavy and the rent of vehicle-mounted laser scanning system and later handling expense are quite expensive. In the third trend, with the planning map of surface features as the blueprint, AutoCAD software is used to draw the plan of surface features, import the plan into a 3DMAX software and draw a 3D model in an equal proportion. The 3D map of surface features established in this way cannot be called as true 3D GIS, because the 3DMAX software platform has not the functions for analyzing and checking the spatial information.

Technology Routine for True 3D Visualization of Surface Features

Based on MapGIS K9, 3DMAX and Daogle software, this paper designs a new technology routine for the production of true 3D visualization of surface features.

Namely, first Daogle Google interceptor is applied to intercept Google map with information of longitude and latitude and topographic map with the information of elevation. And then, GPS gauge is used to measure 4 typical coordinate points of surface features and correct the base map. And then MapGIS K9 software is adopted to produce 2D sector base map and edit it by entering the traits of each surface feature such as name, usage, and elevation and conduct the texture production and the production of a true 3D model of surface features. Finally, the

Fig. 1 Technology routine for true 3D visualization of surface features

Beginning

Daogle Google interceptor software is applied to acquire the base map

Grid data correction

Vectorization

Production of texture

Model production

Production of MAPGIS model

Production of 3DMAX model

Integration of MAPGIS model and 3DMAX model

Route roaming

Completed
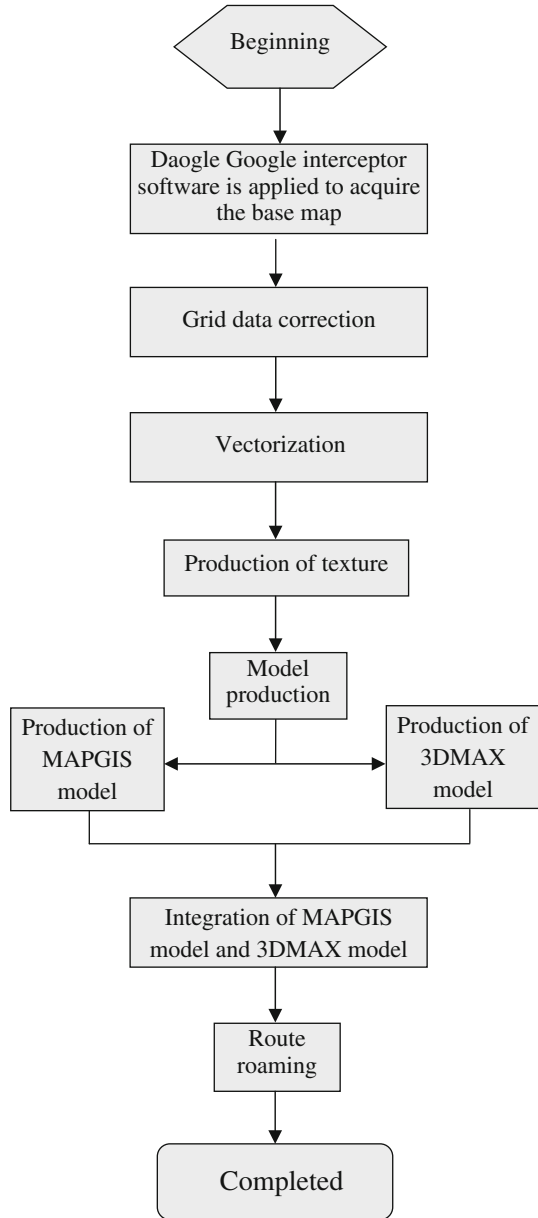
3DMAX software is adopted for the production of other models. And then, the 3DMAX model and MapGIS K9 3D scene is integrated to get the true 3D visualization of surface features. This method has solved the heavy workload in adopting a total station +RTK to measure and get the base map of campus and high

cost of ARCGIS software, and also solved the heavy workload and high cost in the rent of a laser scanner + GPS position finder, which has truly solved the coordinate problem of true 3D surface features (x,y,z) with high efficiency and low cost. The technology routine is shown in Fig. 1.

Production of True 3D Virtual Scene of Surface Features

1. Pre-treatment
(1) Image Data Treatment

The image downloaded with "Daogle Google Map interceptor" is imported into the database. We can see that the image is red and the texture of the architectures is not so clear. Therefore, we need to treat it in the "remote sensing treatment system," choose the "color setting" and re-set the "R,G,B wave bands" to change the color of the image.

(2) Texture Data Treatment

The object of surface features is photographed to collect texture. PS software is mainly used for image cut, rotation, and modification, and finally reaches a real effect of surface features.

For example, in creating the transparent texture, the final display effect of the fencing needs to be hollowed out, which should be displayed to be hollowed out from different perspectives. Here, we will show how to display the texture of a 2D photo into the 3D hollowed out effect.

Get a sample picture of the fence, and then use Photoshop software to extract the contents from the picture for the texture creation. Remove the white background in the picture and only retain the contour of the fence. In the output of the picture, save the picture in the format of PNG, so it can remove the white background. If it is saved into other common formats such as JPEG, the background will be filled with white. In the 3D landscape platform, edit the texture, newly build the texture, and set the "alpha testing" information of the texture parameters. (Testing function): passed if larger or equal to. (Reference value): 100. After saving, the texture is then transparent texture (Fig. 2).

(3) Coordinate Correction

The base map downloaded with "Daogle Google Map interceptor" is of longitude and latitude. The coordinates of the picture intercepted with Daogle Google Map and the coordinates of Google Map are completely consistent in the preciseness. Therefore, its preciseness completely relies upon the preciseness of

Fig. 2 Design sketch of the fence created with transparent texture
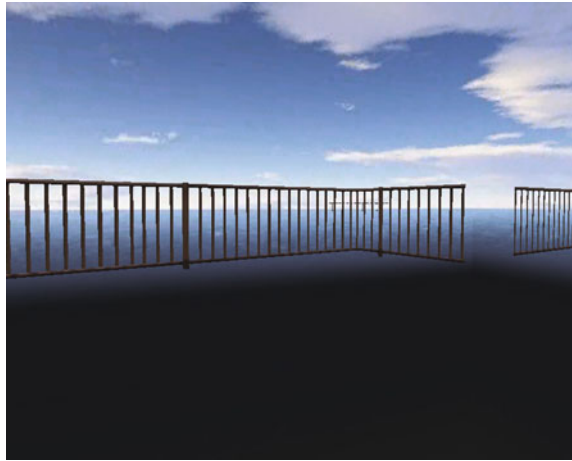
Table 1 Attribute table of line layer

Serial No.	Name	mp Length	mp Layer	Elevation	Color	Line Type
1	Road	676.255354	0	0.01	88	Fold line
15	Trees	589.595777	0	5.00	261	Fold line
18	Architecture	29.911834	0	0.00	432	Fold line
1	Artificial lake	234.35889	0	0.00	261	Smooth line
7	Fencing	1120.3920	0	3.00	585	Fold line
10	Ground track field	116.79024	0	None	733	Fold line

Google Map. Since Google Map has the conditions such as coordinate encryption and bias. Daogle Google Map also has coordinate bias and a little bit deviation. This deviation is basically around 500 m. Therefore, the "base map" will be conducted with grid correction (Table 1), (Fig. 3).

2. Layer Vectorization

The layer of MapGIS is a layer formed from the combination of some related objects by the user in accordance with certain needs or standards. Generally, it is divided into four types of layers, namely, point, line, face, and note. The line layer can also be divided into fencing layer, road layer, tree layer, ground track field layer, etc. As such, in editing the road lines, we can call in the road layer only and shield the layers unrelated with the data of roads, which can greatly reduce the layer data and avoid the interference of unrelated graphics, highlight the parts to be displayed and can also enhance the display speed of the screen (Fig. 4), (Table 2).

In the 3D architecture platform, the roads, sidewalks, trees, and fencing are all established through "line files." Therefore, we put them in the line layer; we put the
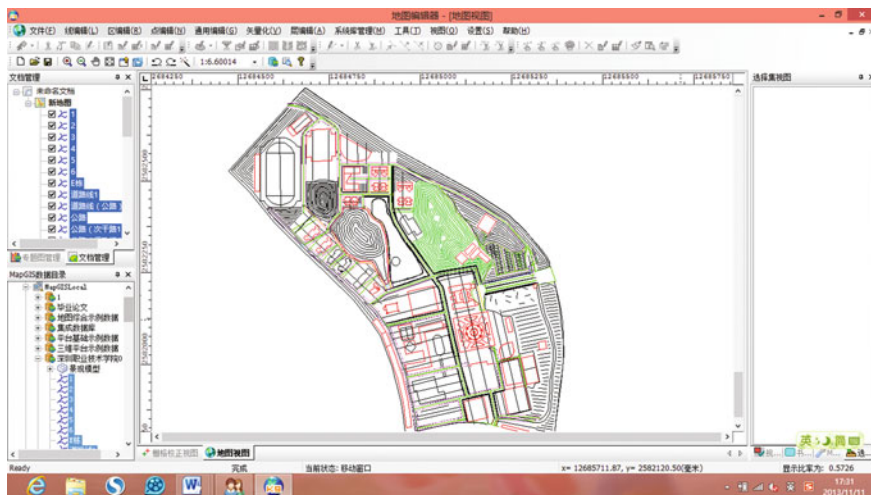
Fig. 3 Line layer
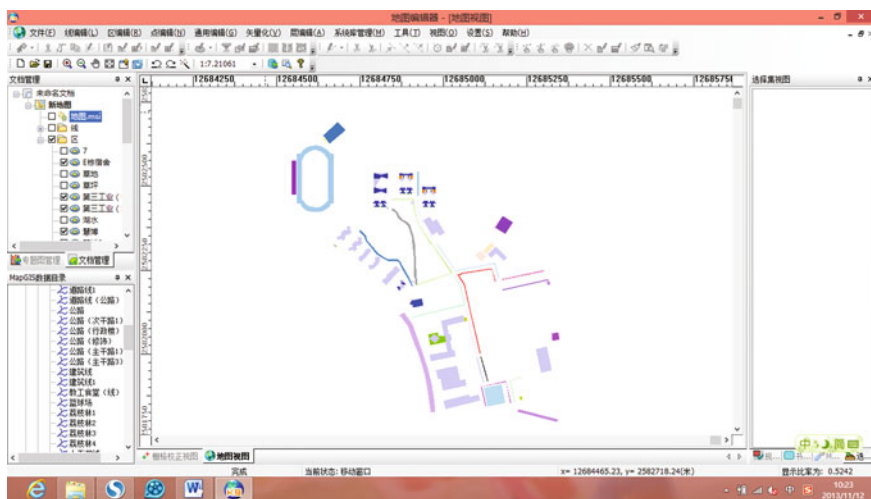
Fig. 4 Area layer

architectures, artificial lake and ground track field in the line layer, because we can establish the area layer through "creating an area by converting line into arc."

3. Modeling of Symbol Surface Features

The architectures such as the teaching building, swimming pool, dorm area, South Gate canteen, artificial lake, litchi woods, roads, Dayun square, the No. 1 industrial center, No. 2 industrial center, No. 3 industrial center, and football

Table 2 Attribute table of area layer

Serial no.	Name	mp Area	mp Perimeter	Elevation	Floors	Color	Mode of filling
1	Lawn	707745.45	437.021877	4.98	None	23	Conventional filling
1	Teaching building	3610.5887	51.497148	25.00	8	96	Conventional filling
1	Faculty's canteen	726.63081	116.310591	16	5	17	Conventional filling
1	Lake water	28983.87	694.87902	4.00	None	49	Conventional filling
12	Dorm (window side)	410.3064	205.73745	21.00	7	279	Conventional filling
11	Dorm (wall side)	508.8441	188.58729	21.00	7	18	Conventional filling

field are the principal and basic components of the campus landscape. Their landscape model shows symbol visual effect, while the appearance of the architectures is relatively regular. The adoption of MapGIS K9 modeling has a real and detailed effect.

This paper is based on the projection numbered EPSG:3785 of Google MAP. This kind of projection is similar to the coordinate system of Web Mecator projection. Its major difference from the conventional Mercator projection is the earth is simulated into a globe rather than a spheroid, which is an equal-angle projection. First, vectorize the base map of surface features acquired from Google Earth. After the vectorization is finished, newly build a line layer. Draw the contour of architectures in the vectorized base map. And then convert it into area layer through the "shifting to area by converting line to arc" and set the elevation attribute and name of the layer. And then, open the MapGIS K9 3D platform software, choose the relevant 2D database and newly build a 3D database. In the following, click the (Landscape Modeling Tool), choose to establish the texture architecture, open the architecture layer, and enter the elevation data into the attribute of high floors of the architectures as the stretch field. Before the choice, use the texture created with the texture tool to achieve automatic 3D modeling of the architectures (Figs. 5 and 6).

4. Modeling of Irregular Surface Features

The imitation model of some irregular surface features in the target area such as road lamps, stools, steles, sculptures, architectures with an arch top and fitness equipment can greatly enhance and enrich the sense of reality and layer of the true 3D visualized campus scene; the 3DMAX software can create very delicate

Fig. 5 Sketch of trees
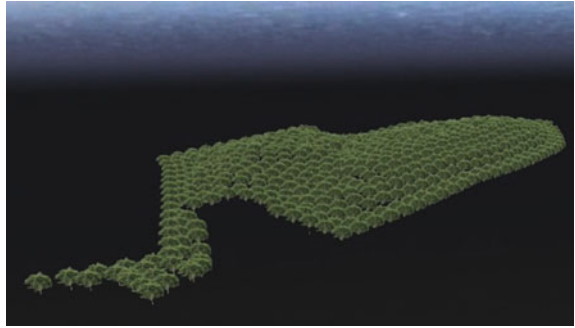
Fig. 6 Sketch of architectures
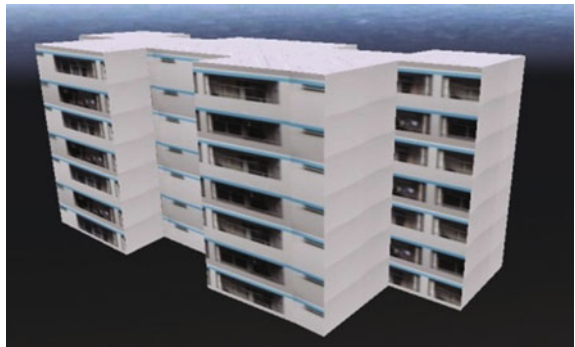
surface features. Therefore, this part will adopt 3DMAX for the modeling. In MapGIS K9, we can import the external 3D model in the format of 3DStudio MAX (*.3ds). However, what is worth attention is the 3D model might be inconsistent with the actual surface features in the direction and size. After downloading, we need to make some adjustment (Figs. 7 and 8).

Taking the badminton gym as an example, its top and facade are arc, which cannot be modeled with MapGIS K9, but only can be modeled with 3DMAX. The model is then imported into the MapGIS K9 3D platform (Figs. 9 and 10).

5. Modeling of Surface Features on 3D Terrain

The research object of this paper is the campus. To truly and objectively reflect the true 3D virtual campus and achieve the function of spatial analysis, we need to overlay the 3D surface features on the basis of elevation model (DEM). For example, in studying how to establish trees on the DEM surface, first we should use the isoheight elevation data H_i of the Google Earth topographic map intercepted with the Daogle interceptor, the measured average height of litchi trees is L_i, the isoheight elevation data of the slopes is $Z_i = H_i–L_i$. Since the same isoheight line is

Fig. 7 Sketch of flowing lake water
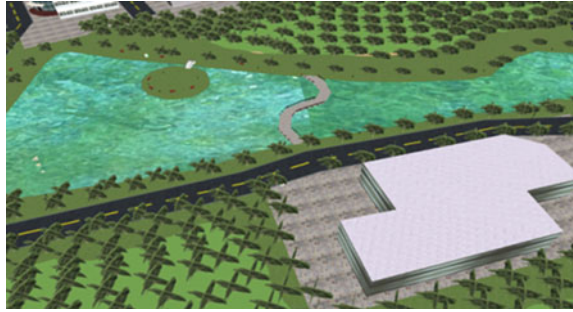
Fig. 8 Sketch of roads
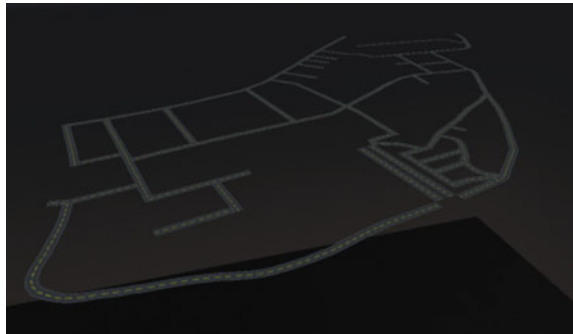
consistent theoretically, we can use an elevation value to express it. Z value is as Formula (1). And through "gridding of discrete data," we will acquire the DEM of the slopes (Table 3).

$$Z = \frac{1}{n}\sum_{i=1}^{n} Z_i \tag{1}$$

And then, we will establish the trees based on the DEM surface, as shown in Fig. 11.

By comparing Figs. 11 and 12, we can find:

Therefore, in establishing trees on DEM surface, we should establish different "line layers" targeting different isoheight lines, and then endow each layer with different elevation values. Meanwhile, the difference distance of trees should also be changed too.

6. Integrated Display of Campus Scene
(1) Integration of architecture models.

It is relatively easy to establish 3D architecture models with MapGIS, but there are many limitations. For example, the flashing board in front of the teaching building, it has a hollowed out design at the surface, which cannot be established

Fig. 9 Sketch of the facade

Fig. 10 Sketch of the right side
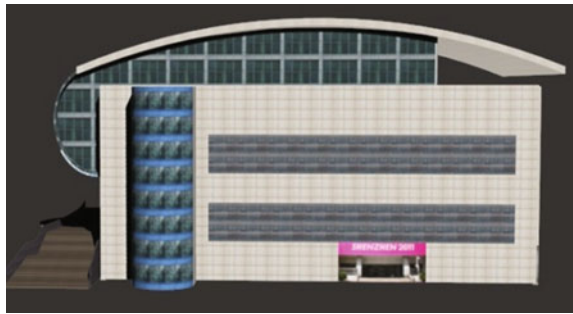
Table 3 Comparison of Attributes

	Trees at the surface of DEM	Trees at the ground surface
Difference distance	Different	Same
Elevation	Different	Same
Height	Same	Same
Width	Same	Same

with MapGIS. Therefore, this part needs to be established with 3DMAX. If the whole architecture model is established with 3DMAX software, the process will be relatively complex. Therefore, finally we use the two kinds of software in combination to establish the architecture model of "teaching building."

The key step is keyboard positioning. Open the "keyboard positioning" in MapGIS K9, add the 3DMAX file, and find the longitude and latitude coordinates of the surface feature to be entered on the vector picture and enter it, set the zoom ratio, thereby making the two architecture models with the same scale and coordinate and integrated with each other.

As shown in Fig. 13, the major architecture of teaching building establishes the model with MapGIS, because the architecture structure is relatively simple; the part indicated with red box will establish the model with 3DMAX software, for the

Fig. 11 Trees at the surface of DEM

Fig. 12 Trees at the ground surface

architecture structure of this part is relatively complex, with hollowed out and round column design at the middle.

(2) Scene Integration

The main body of the visualized campus to be produced with MapGIS K9 3D platform includes various architecture models, 3D topography, lakes, and trees, and so on. And then integrate the irregular shape model produced with 3DMAX into the main body of the visualized campus produced with MapGIS K9 3D platform, which can bring into a full play the advantages of two kinds of software, thereby making the true 3D visualized campus more vivid and delicate. The key for integration lies in confirming the right size of 3DMAX model in accordance with MapGIS K9 3D model to correctly position the geographic coordinates of the area

Fig. 13 Integration sketch of 3DMAX architectures and MapGIS K9

Fig. 14 Overall bird's eye view

to be integrated, thereby making the 3DMAX model to find out the position quickly on the map. The sketch after the integration is shown as Fig. 14.

Spatial Analysis of True 3D Visualized Campus

Actually the true 3D GIS pays most attention to two key points. One is the model expression of true 3D space, the other is the spatial analysis and inquiry of true 3D GIS. MapGIS K9 provides rich spatial analysis and inquiry functions for true 3D GIS, such as single-point topography parameter inquiry, slope analysis, slope aspect analysis, topography surface distance measurement, two-point inter-visibility judgment, flood inundation demonstration, analysis of visual field, calculation of filling and excavation, topography sectioning and sunlight analysis. As shown in Fig. 15, the length of roads in the campus in the true 3D scene is measured. As shown in Fig. 16 is the attribute inquiry of the architectures in the true 3D scene.

Fig. 15 Measurement of ground surface distance

Fig. 16 Attribute inquiry of the architectures

Roaming in the True 3D Visualized Campus

The roaming in the true 3D visualized campus means that the user can observe any surface feature in the space from any direction in the life-like true 3D visual virtual environment and get a feeling as if personally on the scene. In the true 3D

visualized campus, the roaming can be fulfilled with a 3D platform software. Open the 3D database, choose the path roaming tool, and open the interface of path edition. Edit the path in the 3D map, and then preview the path, and then choose the video record in the output tool and finally save the video.

Conclusion

The true 3D visualized campus is a part of constructing the digital city. Constructing a true 3D visualized campus is an experiment for true 3D digital urban construction. This paper takes Xili Lake Campus of Shenzhen Polytechnic as an example, targets the common problems of 3D geometric data such as heavy workload and high modeling costs, analyzes and studies the 2D information and topography information of Google Map acquired with Daogle interceptor and greatly reduces the workload for acquiring the 3D data; uses the MapGIS K9 software to establish a basic 3D scene and finish the integration with the models produced with 3DMAX software, which can effectively reduce the modeling cost; finally it conducts spatial analysis and inquiry on the true 3D visualized campus and achieves the roaming in the true 3D visualized scene. The true 3D visualized campus established with this method features a low cost, a high efficiency, and a good effect.

References

1. Wu, L., W. Shi, and C. Gold. 2003. Spatial modeling technology in 3D GIS and 3D GMS. *Geography and Geographic Information Science* 19 (1): 5–11.
2. He, J., B. Zhang, and Z. Sun. 2013. Production and release of virtual scene based on GIS. *Science of Surveying and Mapping* 38 (4): 172–174.
3. Google Earth Image Download Software–Daogle Google Map Interceptor-V1.1-eiafans Software and Tool Communication Area http://www.eiafans.com/forum.php?mod=viewthread&tid=333462.
4. Xia Chunlin, Wang Jiaqi. 2011. Overview on architecture 3D modeling technology in 3DGIS. *Science of Surveying and Mapping* 36(1): 70–72.

5. Lu, Y., M. Di Xu, and X. Zhang. 2012. Design and realization of 3D virtual campus based on GIS. *Academic Journal of Nanjing Institute of Technology* 4 (1): 81–86.
6. Mu, Y., J. Liu, and Y. Zhang. 2012. Design and realization of 3D campus system based on ArcGIS engine. *Aeronautical Computing Technology* 39 (6): 100–104.
7. Wu, Y., Z. Hu, Y. Huang, L. Lin, and J. Wang. 2010. Design and realization of 3D virtual campus of xuzhou normal university. *Academic Journal of Xuzhou Normal University (Edition of Natural Science)* 28 (1): 75–78.
8. Xu, H., X. Fan, J. Ren, and Z. Zhang. 2010. Urban 3D visualization research based on SketchUp and ArcGIS. *Bulletin of Surveying and Mapping* 3: 52–54.
9. Tang, Z., X. Zhang, and Cao Kaibin. 2010. Research on the application of 3D technology based on skyline in urban planning. *Bulletin of Surveying and Mapping* 5: 10–12.
10. Wang, A., and J. Zhang. 2012. Matching method for physical model and topographic model of 3D GIS architectures and roads. *Bulletin of Surveying and Mapping* 6: 84–89.

'Discontinuity Effect' of Edge Line Markings on Time Headway in Car-Following

Naikan Ding, Shunying Zhu, Hong Wang and Nisha Jiao

Abstract To adjust the time headway 'beforehand' for collision avoidance, the impacts of a kind of designed edge line markings on time headways (THW) were studied. On-road experiment and field observations were conducted to obtain naturalistic vehicle flow data. Statistical analyses showed that (1) the THW increased in the conditions of edge line markings with horizontal offset (HOS) or longitudinal gap (LG) and (2) the THW increased along with the increase of HOS and with the decrease of LG. From the perspective of speed perception and distance perception, the results were interpreted as (1) the edge rate (ER) resulted in speed overestimation, which led to a reduction in actual speed, and (2) the 'discontinuity effect' from edge line markings with HOS and/or LG led to distance underestimation. For safety, drivers incline to increase the distance headways with results of the increase in THW. This may be a new method for preventing car-following accidents.

Keywords Car-following · Time headway · Edge line markings · Visual perception · On-road experiment

N. Ding (✉) · S. Zhu · H. Wang
School of Transportation, Wuhan University of Technology,
1178, Heping Avenue, Wuhan 430063, Hubei, China
e-mail: andrei8901@gmail.com

S. Zhu
e-mail: zhusy2001@163.com

H. Wang
e-mail: wanghong2004317@sina.com

N. Jiao
Department of Transportation of Hubei Province, Planning Research Studio,
428, Jianshe Avenue, Wuhan 430030, Hubei, China
e-mail: 421946755@qq.com

Introduction

Rear-end crashes, accounted for 32.2% of all road crashes, lead to 1694 death and 476,000 injuries in the USA in 2010 [1]. The situation was worse even in China, where 2691 people died because of highway rear-end crashes which account for 40.4% of all highway crashes in 2010 [2]. Researches have shown that various reasons, such as misoperation, slacked reaction, overspeed, and small headway, lead to rear-end crashes. Accordingly, methods of transportation planning, design and policy are extensively used in dealing with road safety issues like the rear-end crashes. However, the yet frequently happened rear-end crashes cast a shadow on these endeavors. Moreover, we found that a great deal of existing methods is more prone to deal with this problem in a way of 'after the event.' However, 'after the event' commonly means that the accidents were already happened and probably with casualty and economic loss. Hence, new beforehand preventing approaches are urgently in need.

From this perspective, in this paper, we naturally focus our attention on the drivers, which is the only element can behave subjectively in the traffic system. It is reported that 90% of all accidents relate to driver's errors or mistakes [3] and 90% information that drivers use is visual [4]. Specifically, for safety, it is of vital importance for drivers to judge the self-speed and distance in car-following. Thus, in order to adjust the drivers' behavior 'beforehand' in car-following, we studied the effects of edge line markings on time headways (THW) from a viewpoint of speed perception and distance perception.

Literature Review

Headways (both time headway and distance headway) have been extensively studied as an essential indicator for evaluating car-following safety [5–9]. Time headway is the time interval between two vehicles in car-following, calculated as distance headway, bumper to bumper distance between two consecutive vehicles in a lane, divided by the speed of the following vehicle. Apparently, the THW in car-following can be influenced by the judgment of speed and distance from drivers.

Actually, from the perspective of visual perception, the estimations or judgments of speed and distance are greatly impacted by the visual information on ground surface. Edge rate (ER) is one of this visual information that has been used in speed control, which is defined as the rate at which local discontinuities cross a fixed point of reference in the observer's field of view [10]. According to this definition, the ER can generally be measured in unit of Hz. Based on virtual-reality experiments, Francois et al. [11] predicted that ER, generated by textures on the ground surface, could lead to the overestimation of self-speed of drivers, which in turn leads to a speed reduction. Larish and Flach [12] discovered that the estimated speed rose

along the increase of ER in its research of driver's estimation of speed of self-motion, with ER, global optic flow rate and texture form regarded as independent variables. Rakha et al. [13] designed a form of transverse bars with a constant gap, and the analyses of data from field observations show a 6 km/h reduction on mean speed and an 8 km/h reduction on 85% speed. Furthermore, by simulation, Liu et al. [14] found that the drivers experienced the maximum speed overestimation when ER of edge line markings varied from 8 to 16 Hz, which also led to a maximum reduction in speed. However, few studies in the above considered the effects of visual information on THW. Thus, the change of THW under the circumstance of speed reduction from edge line markings is needed to be examined.

On the other hand, the perception of distance is also affected by visual information like the surface textures. Gibson [15] puts forward the 'Ground Theory,' and it predicted that when the common ground surface is disrupted, the visual system is unable to establish a reliable reference frame and consequently fails to obtain correct absolute distance. Sinai et al. [16] used 'the blindfolded walking task' and 'the perceptual matching task' in its experiments, and the results showed that distance underestimation occurred in the condition of texture discontinuities, which to certain extent verifies the prediction of Gibson. Similarly, Wu et al. [17] and Yarbrough et al. [18] presented discontinuous textures on the ground in virtual-reality environment and found that the distance was underestimated by observers. Besides, Feria et al. [19] demonstrated offset stripes in varying widths and gaps in a simulated scenario, and an underestimation in distance was also observed. Additionally, Feria referred this phenomenon as the 'discontinuity effect.' However, these experiments about distance perception were primarily conducted in static virtual-reality scenes, which were quite different from that in situations of driving in real world. Thus, it is needed to inspect if there is a similar effect of distance underestimation from edge line markings on drivers in car-following.

In fact, it could be concluded that the discontinuity of texture (like the markings) on the ground surface is perhaps the key factor that give rise to the underestimation in distance and even the overestimation in speed. It means the 'discontinuity effect' could be generated from edge line markings, with certain artificial designs. More importantly, 'discontinuity' can simply be classified to horizontal offset (HOS) and longitudinal gap (LG) of strips (in Feria's experiment) or markings (in this study). Actually, the LG of strips or markings has already been demonstrated in those studies regarding ER. Because, according to the definition, ER is the rate of discontinuities cross a fixed point of reference. It means the speed overestimation of a driver could be caused by edge line markings with longitudinal gap. In addition, distance underestimation could also be produced by edge line markings with both longitudinal gap and horizontal offset for discontinuities exist. What's more, the overestimation in speed and underestimation in distance may lead to changes in time headway in car-following.

Based on the above thoughts, the research reported here aims at examining the 'discontinuity effect' of edge line markings on time headway in car-following. For this purpose, two kinds of simplified discontinuous edge line markings were designed, that is, line markings with HOS and line markings with LG.

The data used for statistical analysis of headways came from on-road experiment and field observations carried out on an expressway in Wuhan, Hubei Province, China.

This paper is organized as follows. The next section gives a brief introduction of the methods we use and the experiment. In the following section, the effects of HOS and LG of edge line markings were analyzed and demonstrated. And this is followed by a discussion of the THW from the perspective of speed perception and distance perception in the fourth section. The final section concludes this research effort and proposes further research work.

Method and Experiment

Methods Overview

In this study, the HOS and LG of edge line markings were taken as independent variables to investigate its effects on THW in car-following. Three levels were set up for both HOS and LG, which will be detailed later in this part. In order to test the possible 'discontinuous effect' from edge line markings in real scene, the methods of on-road experiment and field observation were employed to obtain naturalistic car-following data. Furthermore, to exhibit edge line markings on road surface, a kind of yellow pre-formed adhesive tape was installed aside of the two original white markings of the slow lane. In addition, the influence of the markings per se was analyzed with the original situation regarded as a control. For these purpose, data pairs of speed and time headways of every single vehicle at consecutive observation sections were collected.

Test Site

The segment of Daijiashan and Huangpi Expressway (coded S1) at Wuhan, Hubei Province, P.R. China was selected as test site. It was a two-way-four-lane expressway and the design speed was 100 km/h and land width was 3.75 m. The Annual Average Daily Traffic (AADT) was 18,061 veh/day, and the percentage of heavy vehicles was 7.8%. The precise location of this experimental site was between K6 + 900 and K8 + 000, at which it was a flat and straight segment. Along the direction of mileage increment, the straight segment was connected with a flat and straight segment of 900 m in length in upstream and with a right-turn curve in downstream. Besides, in the middle part of K6 + 900 to K8 + 000, a length of 300 m slow lane was installed with edge line markings. There was no tunnel or overpass within the experimental site area, nor any exposed surveillance system for violations capture.

Design of Edge Line Markings

Horizontal Offset (HOS)

In order to make the offset width the only variable, all other parameters of the markings were set unified, which include the length of one single marking (100 cm) and the width of one single marking (15 cm). Besides, from the perspective of design, one critical issue is to determine the minimum width of the offset. For this issue, the visual acuity and distance were taken into consideration. The visual acuity is a measure of the spatial resolution of the visual processing system, which can be calculated as follows: Acuity $= 1/$gap size (arc min) [20]. Here, the gap size is the critical visual angle between the offset width of the markings and the eye of a driver, when the offset of the markings can just be seen from the driver in distance. In fact, there is a simple geographic relation among the visual angle (α), distance (D), and HOS, and that is $HOS/2D = \tan \alpha/2$, $HOS = 2D \cdot \tan \alpha/2$. Notably, dynamic sight should be employed since the drivers move in a relatively high speed. Besides, as the design speed of the test expressway equals 100 km/h, the average visual acuity of drivers is 0.3, that means the corresponding visual angle (α) is $0.00097°$. Also, here the stopping sight distance was regarded as the critical distance (D), which is 160 m when the design speed is 100 km/h. Then, $HOS = 2 \times 160 \times \tan(0.00097/2) = 15.5$ cm. To facilitate the work of installation of the markings, the minimum HOS was set to 20 cm, and other two levels of HOS were set to 30 and 40 cm. Figure 1 shows the horizontal offset design of edge line markings and the actual layout on roadway surface.

Longitudinal Gap (LG)

Similarly, we chose three levels of gap between two adjacent markings. That was, 100, 125, and 150 cm, and the corresponding unit length of the markings are 200, 250, and 300 cm, respectively. Notably, the horizontal offset here was fixed to 20 cm. Figure 2 depicts the longitudinal gap design of edge line markings and the actual layout on roadway surface.

Data Collection and Treatment

In this experiment, the speed and time headways of each vehicle were obtained indirectly by recording the very moments when the vehicle passed through two sections. Figure 3 demonstrates a sketch map of the way the moments collected. As shown in Fig. 3, six video cameras (Sony HDR-PJ510E, 50 frames of images with a resolution of 1920×1080 recorded per second) were consecutively mounted outside of the crash barrier on the hard shoulder with a constant interval length of 100 m
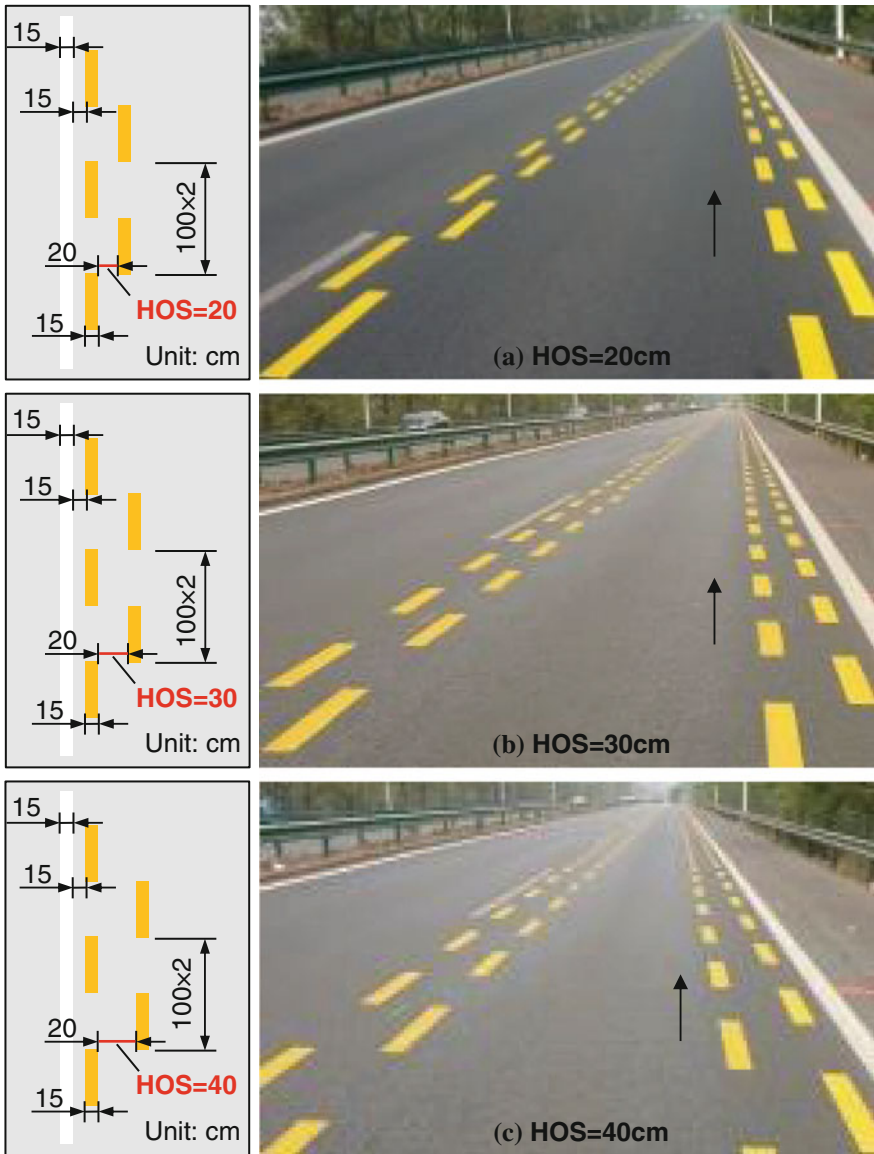
Fig. 1 Designs of edge line markings with Horizontal Offset (HOS). The *left* parts illustrate sketch maps of the designs, and the *right* parts show the real scenes

in depth. Moreover, to avoid the impact from drivers who would take those cameras as traffic violation surveillance and adjust driving behavior, the cameras were sheltered with shrubs and invisible from the lanes. Accordingly, the six video cameras recorded vehicles at six different sections, respectively, which represent six
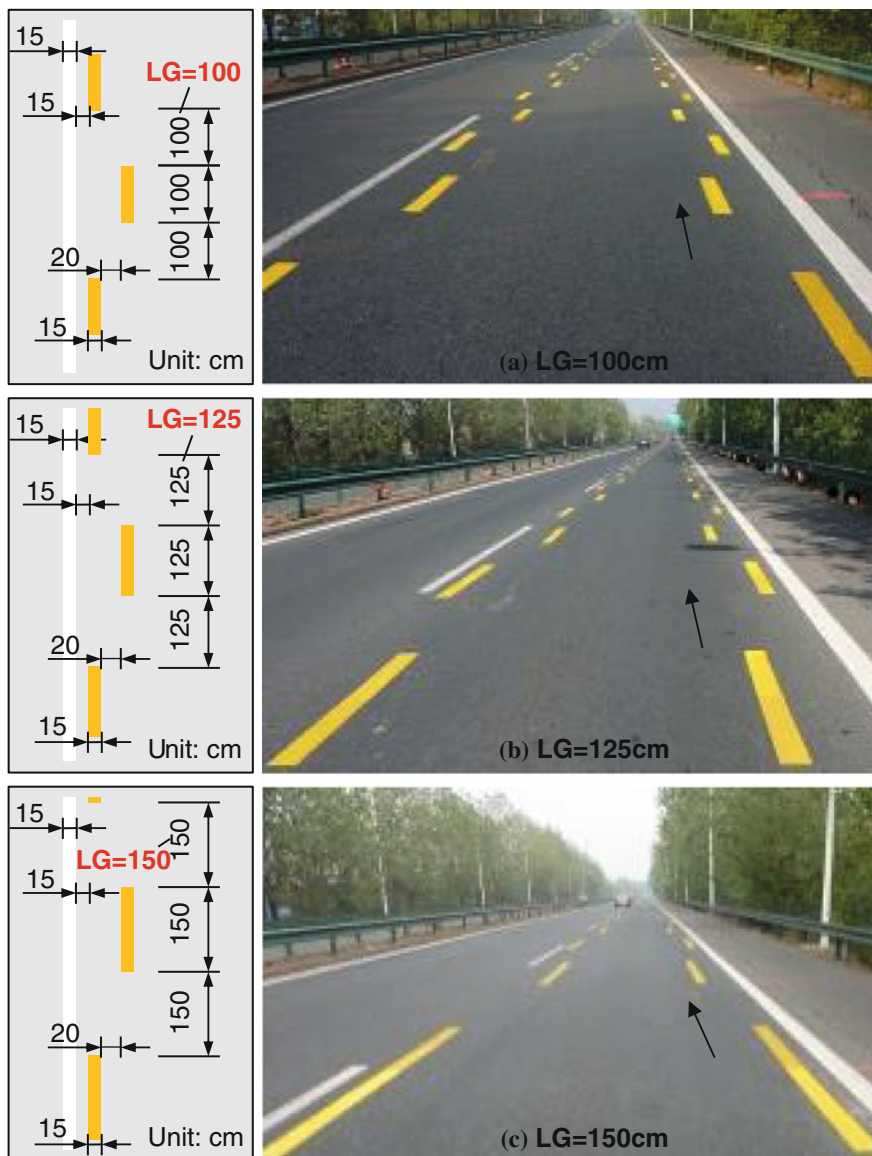
Fig. 2 Designs of edge line markings with Longitudinal Gap (LG). The *left* parts illustrate sketch maps of the designs, and the *right* parts show the real scenes

observation sections (numbered in the bottom of Fig. 3). In addition, in order to minimize the observation error, a total length of 500 m area was artificially scaled like a 'rule' with red tick marks (see Fig. 3). Thus, the speed and time headway of vehicles at section 1–6 could be calculated as follows. Take section 2 as an example,
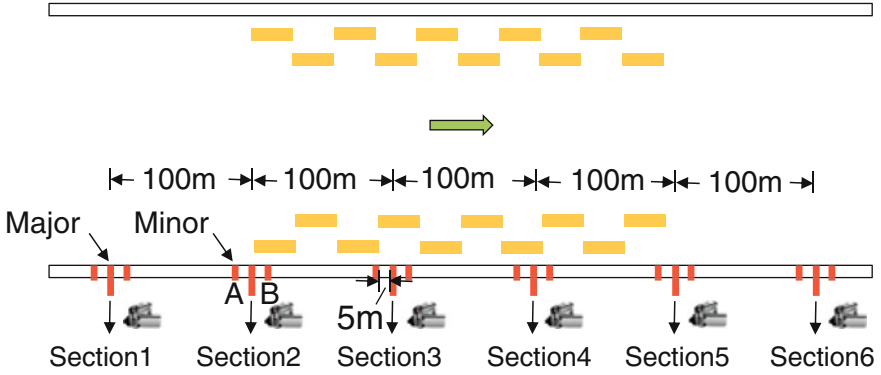
Fig. 3 Layout of the cameras

$v_2 = (x_B - x_A)/(t_B - t_A)$, where v_2 is the speed of a vehicle when it just pass through section 2, x_A and x_B are the positions of the vehicle at minor section A and B, t_A and t_B are the time (or moment) of the vehicle at minor section A and B. In particular, in this experiment the minor section A and B were set 5 m away from its corresponding major section. It means $x_B - x_A = 10$ m. Similar calculations were conducted in the other five major sections. Moreover, the time (moment) that vehicles passed through the major section 1–6 was also recorded, so the time headway for a single vehicle can be calculated as $THW_i^k = t_{i+1}^k - t_i^k$, where THW_i^k is the time headway when vehicle i passes through the major section k, t_i^k and t_{i+1}^k are the very time when vehicle i (leading vehicle) and vehicle $i+1$ (following vehicle) pass through the major section k. The average speed and/or time headway of the section 2–5 was used for comparisons among effects of the three levels of HOS and LG. The average speed and/or time headway of every single major section was used to check the possible difference among sections. All the data were collected at 8:30–11:30 a.m. or 14:00–17:00 p.m. in no precipitation days. Besides, in order to meet the request of sample size for statistical analysis, at least one day observation was conducted in every single test.

Data Filtering Process

Filter Out Free-Flow Vehicles

A state when there is no constraint placed on a driver by other vehicles on the road is called free-flow state. To guarantee the data validity used for car-following analysis, those free-flow vehicles were filtered out in such a criterion: If the stopping time of a vehicle was less than its time headway, then the vehicle should be defined as free-flow vehicle, otherwise, it was a vehicle in car-following.

Here, the stopping time was calculated as follows: $t = V/a$, where t is the stopping time, s; V is the instantaneous speed of a vehicle, m/s; a is the deceleration ($a = 2.5 \text{ m/s}^2$ was suggested by AASHTO [21]). In addition, if the state of free-flow occurred at any section of the six, then the vehicle needed to be removed.

Filter Out Lane-Change Vehicles

There was the possibility that lane-change would happen of vehicles. Thus, to avoid the influence of these data, video images were employed. The trajectory of each vehicle from the 6 video clips was reviewed frame by frame to check if there was lane-change vehicle. If lane-change was found between any two consecutive observation sections, then this vehicle needed to be removed from the dataset.

Based on the data collecting methodologies and data filtering rules, the effective sample of each test was showed as follows in Table 1.

Results

Effects of HOS

Figure 4a shows the THW (mean ± s.e.m.) in each test. It can be discovered that there were increases of THW of HOS = 20 cm (0.02 s or 0.5%), HOS = 30 cm (0.10 s or 2.4%), and HOS = 40 cm (0.15 s or 3.6%) when compared with the control test where no extra markings were installed. Figure 4b shows the THW of each section among the three tests. The relation of THW of the three tests (i.e., HOS = 20 < HOS = 30 < HOS = 40 cm) was found to be consistent among the six sections. Additionally, it can be found that the average time headway increased a little bit through section 1–6 in the test of every levels of HOS.

Table 1 Effective sample size of each test

Test		Sample					
		Original			Effective		
		Morning	Afternoon	Total	Morning	Afternoon	Total
Control		1387	816	2203	289	203	492
HOS (cm)	20	1755	1108	2863	327	429	756
	30	1108	1057	2165	246	357	603
	40	801	949	1750	169	123	292
LG (cm)	100	1113	546	1659	193	132	325
	125	890	348	1238	161	117	278
	150	737	895	1632	224	140	364

(a)

HOS verses THW

(b)

Time Headways of Each Section
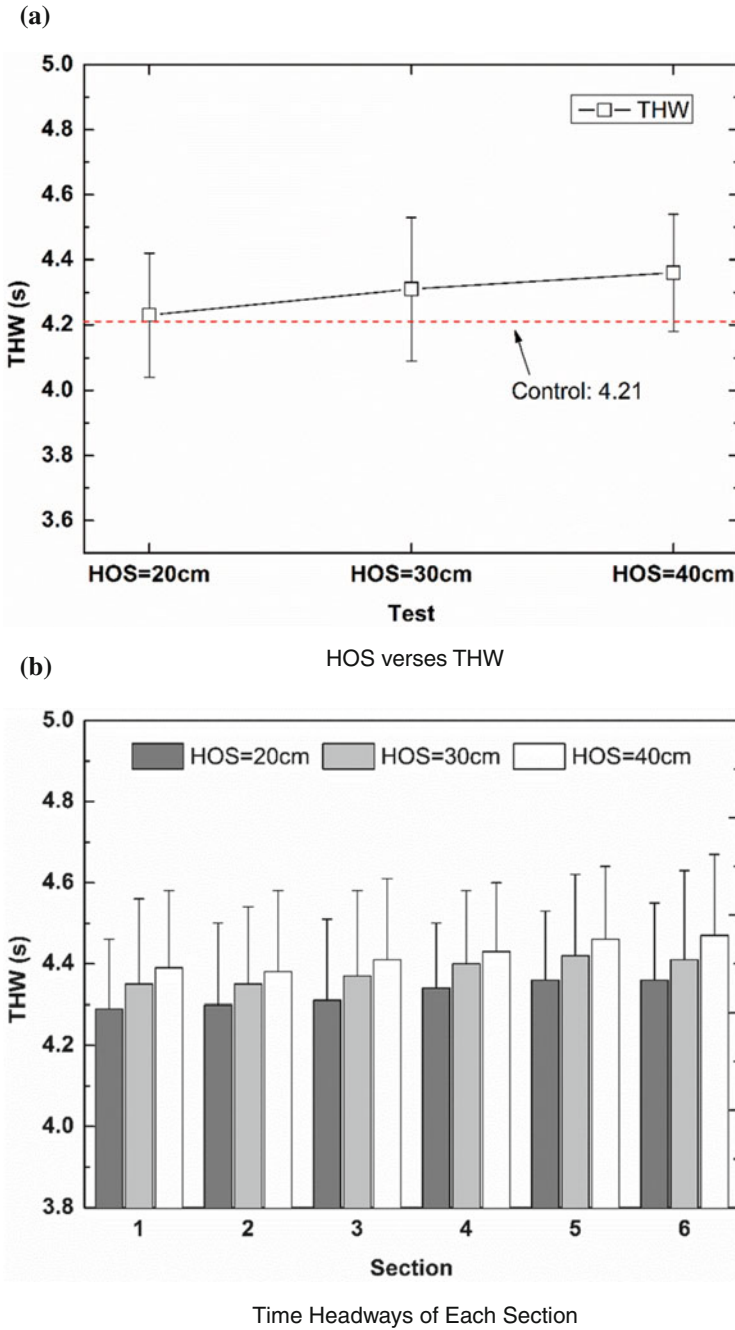
Fig. 4 Mean time headways of HOS tests

Moreover, the effect of HOS on the time headway was found to be significant by one-way ANOVA ($F(2,1648) = 137.69$, $P < 0.05$).

Effects of LG

Figure 5a shows the THW (mean $\pm$ s.e.m.) in each test. It can be found that the THW varied among the three tests of LG, and its increments with regard to the control were also varied in certain extent (LG = 100 cm: 0.22 or 5.2%; LG = 125 cm: 0.16 or 3.8%; LG = 150 cm: 0.13 or 3.1%). Figure 5b shows the THW of each section among the three tests. The relation of THW of the three tests (i.e., LG = 100 > LG = 125 > LG = 150 cm) was found to be consistent among the six sections. Additionally, it can be found that the average time headway increased a little bit through section 1–6 in the test of every levels of LG. Furthermore, the effect of LG on the average time headway was found to be significant by one-way ANOVA ($F(2,964) = 114.37$, $P < 0.05$).

Comparison with the Control Test

In addition to the analyses of effects of HOS and LG, respectively, a comparison of HOS test and LG test with the control test was as well conducted. Figure 6 demonstrates the THW (mean $\pm$ s.e.m.) of these three conditions. By contrast, it was found that the average THW of HOS test and LG test, in which edge line markings were installed, were significantly greater than that of control where there was no extra edge line marking. (Asterisks mark a significant difference (*$P < 0.05$, **$P < 0.01$, unpaired Student's t test)).

Discussions

Results of the present study suggest several conclusions in regard to the effects of visual information on car-following performance. Firstly, for the test of HOS, the average time headway increased as the HOS increased. Secondly, for the test of LG, the average time headway decreased as the LG increased. Meanwhile, within the scope of observations, the time headway rose along with the forward of vehicles on the lane. Besides, it can also be concluded from the result of comparisons with the control test that the existing of the special designed edge line markings per se can impact the THW in car-following. Here, the increment of THW may be attributed to the perception of speed and distance that influenced by edge line markings, since the time headway comprises the factors of speed and distance headway.

(a)

LG verses THW

(b)

Time Headways of Each Section
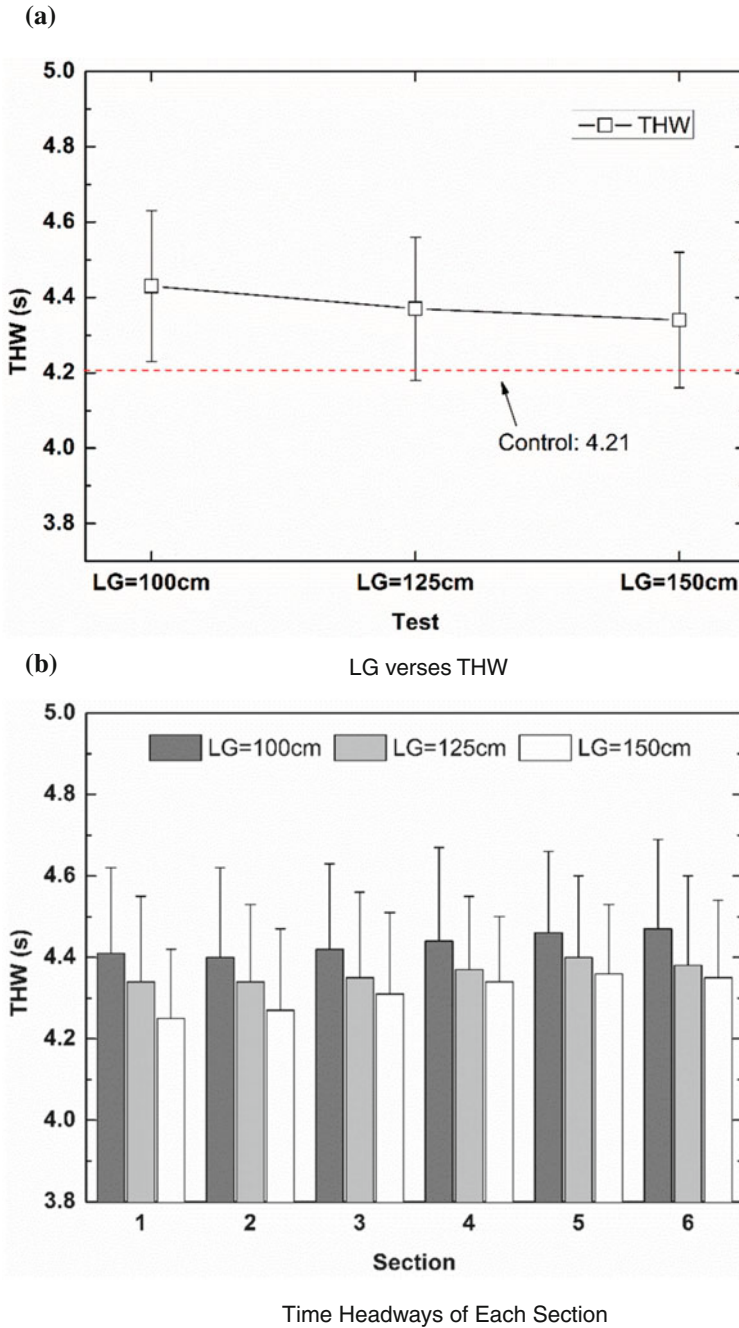
Fig. 5 Mean time headways of LG tests
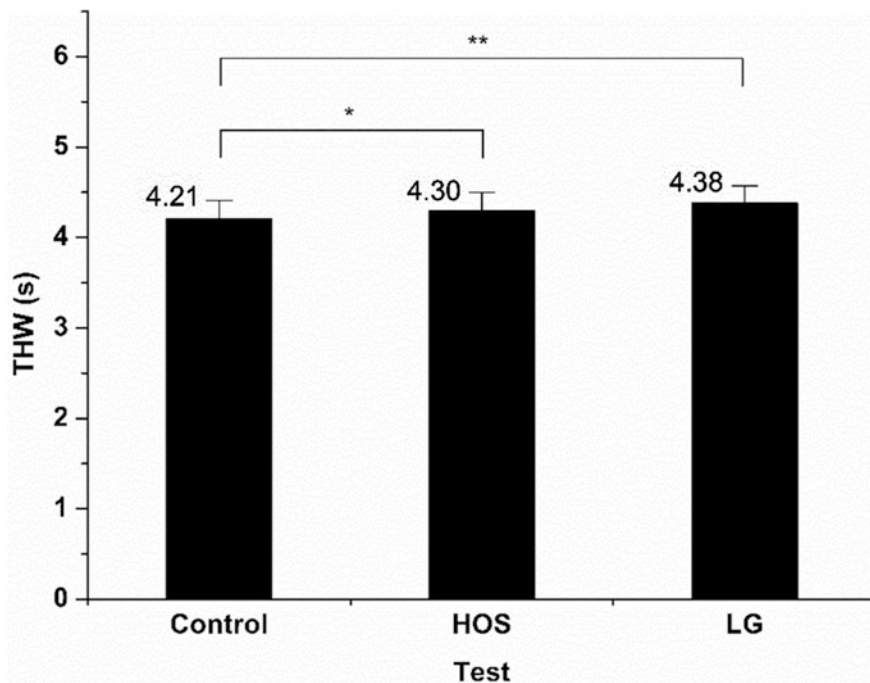
Fig. 6 Mean time headways of the HOS tests, LG tests, and control

Analysis of Speed and Speed Perception

Speed is the key factor that can be controlled by a driver directly. According to the definition of time headway, the increment of THW could be contributed to the reduction of speed. So, an inspection of speed initiated. Figure 7 shows the average speed of all tests. It can be seen that there were speed reductions of HOS tests and LG tests compared with the control test. Obviously, the average speed of LG tests reduced by a larger amount than that of HOS tests with respect to the control. That is, the average speed reduced by 2.33 km/h (or 3.0%) in LG tests, but the counterpart of HOS tests was 0.63 km/h (or 0.8%). Furthermore, a one-way ANOVA showed that there was a significant effect of LG on speed ($F(2,964) = 142.91$, $P < 0.05$), but the effect of HOS on speed was not found to be significant.

Actually, the significant reduction in speed was in line with results of Francis's [11], Rakha's [13], and Liu's [14]. Nevertheless, and more notably, the present study distinguished from these works in the following aspects. There are differences in the experimental sites. A virtual-reality setup was used in Francis's study, but to get the real-world vehicle flow data, the on-road experiment and field observation were conducted in this study. Also, Francis studied the effects of visual information on walking velocity and that was quite different from the situations in car-following on roadway in this study. However, the similar results of studies with different
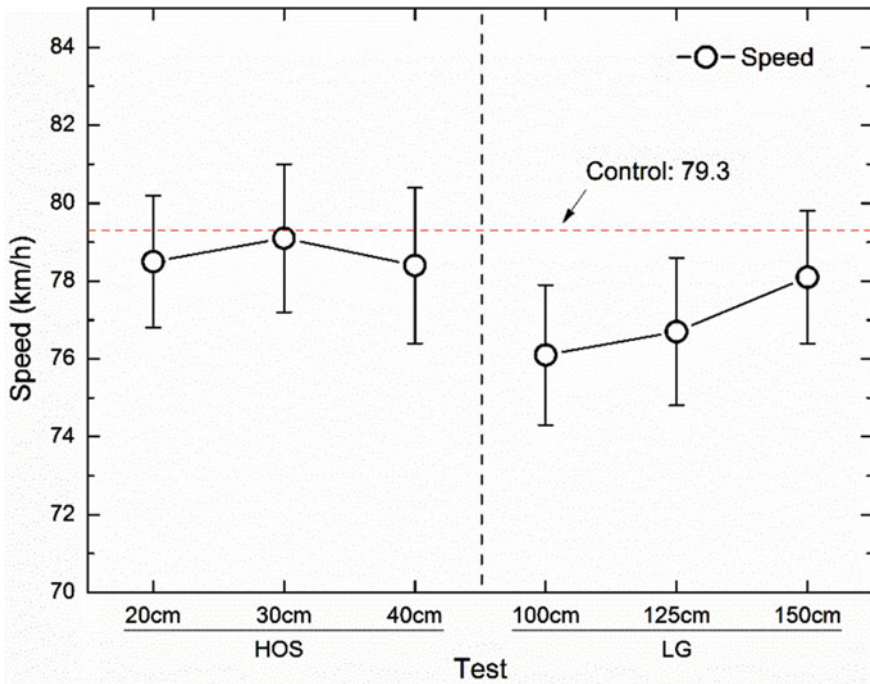
Fig. 7 Speeds of the HOS tests, LG tests, and control test

experimental background seemed to be that something the same here mattered. That was perhaps the ER. According to the definition of ER, a driver will perceive a different speed (self-speed) in conditions of different texture densities on the ground surface and/or different self-speed. Moreover, Francis's and Liu's studies suggested that the increase of the density of edge line markings led to a larger amount of reduction in speed. This was consistent with the results of the test of LG. Thus, ER caused the overestimation in self-speed among drivers, which in turn led to the reductions in self-speed. Consequently, the THW increased.

Analysis of Distance Headway and Distance Perception

Distance is another important factor that influences driving behaviors greatly. Figure 8 presents the average distance headway of all tests. It can be discovered that the average distance headway increased in the HOS tests and LG tests compared with the control. That was the average distance headway increased by 3.83 m (or 4.8%) in HOS tests, and the counterpart of LG tests was 2.1 m (or 2.7%). Furthermore, a one-way ANOVA showed that the effects of HOS and LG on distance headway were both found to be significant (HOS tests: $F(2,1648) = 155.73$, $P < 0.05$; LG tests: $F(2,964) = 138.17$, $P < 0.05$).

As a matter of fact, the density of edge line markings may have influence on the estimation of physical distance (or distance headway) since the 'discontinuity' exists. However, what is firstly noteworthy is that the actual distance needed to be judged in this study was relatively more remote than that in studies of Sinai's [16], Wu's [17], Yarborough's [18], and Feria's [19]. Besides, comparatively more diverse styles of textures were designed in this study, which was obviously distinctive from these previous ones. And this difference in texture design perhaps led to the difference in the perception of actual distance. Additionally, combined with the effects on perceived speed, it can be found that both the leading vehicle and following vehicle were in a state of speed reduction, which was significantly different from the state of low speed. This difference may lead to the difference in the estimation of distance. Nevertheless, a similar 'discontinuity effect' was yet found in a real, open, and high-speed motion scenario. Figure 8 also indicates that a higher density of markings can result in larger distance headways. This relation was consistent with results of the study of Fajen's [22]. The results of Fajen's virtual-reality experiment demonstrated that the average stop distance decreased as the increase of the density of texture (note that here the stop distance was negative as all vehicles exceeded the 'STOP' sign). This means that a high density of markings can lead to an increment in distance headway and with result of increment in THW.
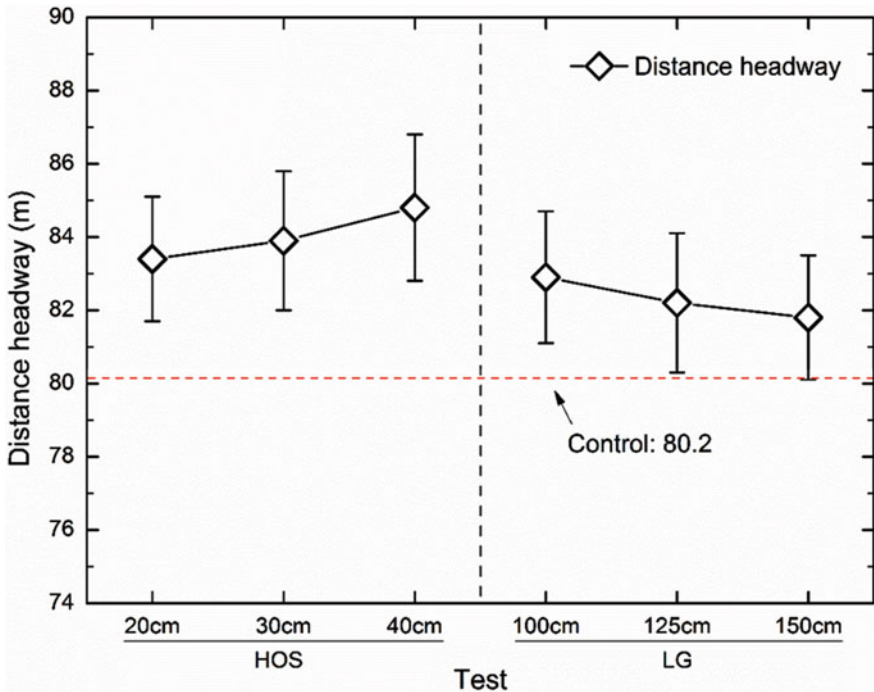
Fig. 8 Distance headways of the HOS tests, LG tests, and control test

Conclusions and Recommendations

The main conclusions from this study are as follows:

1. The existing of edge line markings with HOS and/or LG can lead to the increment of THW when compared with the control in which no extra marking was installed.
2. The effects of edge line markings (with HOS and/or LG) on THW were interpreted with knowledge of speed perception and distance perception. More importantly, reductions in speed and increments in distance headways we found in the condition of installation of the designed edge line markings.
3. There are limitations in our work. First, in this paper, only 300 m was installed with edge line markings (the whole test area was no longer than 500 m), so the effects of edge line markings may have not been produced thoroughly. Secondly, there is yet not enough evidence to support that there could exist a best design of markings with HOS and/or LG to influence THW most effectively. Because only three levels of HOS and LG of edge line markings were adopted.
4. Future works. For the consideration of application, a larger range of the installation area and/or a longer observation time are needed for further verification and analysis. Besides, other important latent variables are needed to be taken into consideration, such as age range, gender, vehicle type (vehicle size), and road alignment.

Additionally, some recommendations could be made for traffic safety by the application of this kind of edge line markings. As the original intention of this study, the application of the designed edge line markings can be used as a 'beforehand' countermeasure for accident prevention at accident black spots. And this 'beforehand' countermeasure could reduce the number of accidents or may decrease the intensity of an accident.

References

1. NHTSA. 2012. Traffic safety facts 2010: A compilation of motor vehicle crash data from the fatality analysis reporting system and the general estimates system. Technical report DOT HS 811 659. NHTSA, U.S. Department of Transportation.
2. Ministry of Public Security Traffic Management Bureau, 2011. White paper on road traffic accidents in China 2010 (in Chinese). Traffic Management Bureau, P.R.C. Ministry of Public Security.
3. Theeuwes, J., and H. Godthelp. 1995. Self-explaining roads. *Safety Science* 19: 217–225.

4. Sivak, M. 1996. The information that drivers use: Is it indeed 90% visual? *Perception* 25 (9): 1081–1089.
5. Risto, M., and M.H. Martens. 2013. Time and space: the difference between following time headway and distance headway instructions. *Transportation Research Part F: Traffic Psychology and Behaviour* 17: 45–51.
6. Habtemichael, F.G., L. de Picado Santos., and N.-E., El Faouzi. 2012. Parameters of time headway distribution as performance indicators of motorway traffic and driver behavior—comparison of good and adverse weather conditions. Transportation Research Record: Journal of the Transportation Research Board, 2272: 152–160.
7. Taieb-Maimon, M. 2007. Learning headway estimation in driving. *Human Factors: The Journal of the Human Factors and Ergonomics Society* 49 (4): 734–744.
8. Lewis-Evans, B., D. De Waard, and K.A. Brookhuis. 2010. That's close enough-a threshold effect of time headway on the experience of risk, task difficulty, effort, and comfort. *Accident Analysis and Prevention* 42 (6): 1926–1933.
9. Lu, G.Q., B. Cheng, Q.F. Lin, and Y.P. Wang. 2012. Quantitative indicator of homeostatic risk perception in car following. *Safety Science* 50 (9): 1898–1905.
10. Warren, R. 1982. Optical transformations during movement: Review of the optical concomitants of egospeed. Final technical report for Grant No. AFOSR-81-0108. Department of Psychology, Ohio State University, Ohio.
11. Francois, M., A.H. Morice, R.J. Bootsma, and G. Montagne. 2011. Visual control of walking velocity. *Neuroscience Research* 70 (2): 214–219.
12. Larish, J.F., and J.M. Flach. 1990. Sources of optical information useful for perception of speed of rectilinear self-motion. *Journal of Experimental Psychology: Human Perception and Performance* 16 (2): 295–302.
13. Rakha, H.A., B.J. Katz., and D. Duke. 2006. Design and evaluation of peripheral transverse bars to reduce vehicle speeds. In *Proceedings of the 85th transportation research board annual meeting*, Washington, D.C.
14. Liu, B., S.Y. Zhu., H. Wang., and L.Z. Cheng. 2012. Optimization design and experiment on plane layout of edge line marking for speed reduction. In *Proceedings of the 92nd Transportation Research Board Annual Meeting*, Washington, D.C.
15. Gibson, J.J. 1950. *The perception of the visual world*. Boston: Houghton Mifflin.
16. Sinai, M.J., T.L. Ooi, and Z.J. He. 1998. Terrain influences the accurate judgment of distance. *Nature* 395 (6701): 497–500.
17. Wu, B., Z.J. He, and T.L. Ooi. 2002. A ground surface based space perception in the virtual environment. *Journal of Vision* 2 (7): 513.
18. Yarbrough, G.L., B. Wu, J. Wu, Z.J. He, and T.L. Ooi. 2002. Judgments of object location behind an obstacle depend on the particular information selected. *Journal of Vision* 2 (7): 625.
19. Feria, C.S., M.L. Braunstein, and G.J. Andersen. 2003. Judging distance across texture discontinuities. *Perception* 32 (12): 1423–1440.
20. Cline, D., H.W. Hofstetter, and J.R. Griffin. 1997. *Dictionary of visual science*, 4th ed. Boston: Butterworth-Heinemann.
21. AASHTO. 2011. A policy on geometric design of highways and streets. 6th ed. USA: AASHTO.
22. Fajen, B.R. 2005. Calibration, information, and control strategies for braking to avoid a collision. *Journal of Experimental Psychology: Human Perception and Performance* 31 (3): 480–501.

Study on the Influence on Liquid Sloshing Caused by Baffle's Parameter Changes in Tank

Hongfei Liu, Hang Lv, Hepeng Wang and Yihua Zhang

Abstract Liquid sloshing in the tank is simulated by numerical method in this thesis. The Volume of Fluid (VOF) method and the standard k-ε turbulent model, and the PISO algorithm is determined for the solution. The thesis investigates numerical simulation of different numbers of the baffles and the baffle location in circle cross-sectional tank with the FLUENT software. The results show that (1) In the circle cross-sectional tank, the effect of defending waves will be better with the increasing number of baffles. (2) The larger the baffle area below the liquid free surface, the better the effect of defending waves.

Keywords Tank · Liquid free surface · Baffle · Liquid sloshing · Impact force

Introduction

Along with the rapid development of Chinese logistics industry, and the dangerous goods traffic continues to increase. Also, the main mode of transport of flammable, explosive, poisonous and other dangerous goods is liquid tank trucks. The stability of tank trucks can become poor due to the shaking of the internal liquid, which will cause accidents easily. And dangerous goods will lead to secondary damage. In order to decrease the influence of liquid sloshing on stability of the tank, baffles are

H. Liu (✉) · H. Lv · Y. Zhang
School of Transportation, Jilin University, Changchun 130022, Jilin, China
e-mail: hongfeiliu@jlu.edu.cn

H. Lv
e-mail: lv_hang_56@163.com

Y. Zhang
e-mail: zhangyihua1989@126.com

H. Wang
Changchun Automobile Industry Institute, Changchun 130011, Jilin, China
e-mail: wanghepeng@gmail.com

used in tanks. The thesis discussed the influence of the number, location and other parameters of baffles on liquid sloshing in tanks with FLUENT software.

There are some studies about this now. Wan Liping, Qian Caifu and others [1] also used volume of fluid (VOF) method to calculate the effects of increase of baffle's number and area in liquid sloshing. And they pointed out that increasing the number of baffles can slow down the changes of liquid impact. But it could not reduce the maximum impact force efficiently. Gao [2] used FLUENT software to calculate the location of baffles. They found when the baffles are put at the top or closely enough to the bottom of tanks, it could have a better effect of defending waves. When the filling ratio was greater than 0.7, the top one could get a better effect, and the bottom one would act well only when the filling ratio was very little. Yue et al. [3] calculated the influence of three typical shapes of baffles on liquid sloshing, and a baffle with relative concentration of the area had a better effect of defending waves. Chen [4] simulated changing the number of baffles, liquid viscosity and other parameters by numerical method when the tank is braking and turning. And he pointed out that in order to prevent rollover, the filling ratio should be greater than 0.8. Cacciatore Philip et al. [5] analysed the stress and stability of the tank when in the typical seismic velocity spectrum using the finite element method. And Cheng et al. [6] researched the elastic diaphragm and liquid coupling vibration in rectangular tank. They get the conclusion that the natural frequency changed with the stiffness and location of elastic diaphragms. Firouz-Abadi and Haddadpour [7] used boundary element method to establish a 3D reduced-order model of liquid sloshing in the tank with their team. They validated the capability and accuracy of reduced-order model when dealing with liquid sloshing problem. According to VOF method, using CFD software to discuss liquid sloshing when the container was partially filled with fluid in rectangular tank, Gradinscak Marija and Jafar Farial established liquid sloshing model. Hou Ling, Li Fangcheng and others also used CFD software and VOF method to simulate liquid sloshing in 2D liquid container by numerical method when it was stimulated by external excitation. They found if tank frequently gets multiple coupling and resonance excitation, then the tank would slosh very sorely and the sloshing load would be more and more.

This thesis carried out a detailed simulation analysis of different number and location of baffles in circle cross-sectional tank and gets the influence of number and area below the liquid free surface on liquid sloshing.

The Influence of the Number of Baffles on Liquid Sloshing in Circle Cross-sectional Tank

We discussed about the influence when the filling ratio is 0.25, 0.5, 0.75 and 0.875.

The design of baffles is complied with the requirement of *Liquefied gas tanker safety rules,* "When installing the baffles, the area of upper bow should be smaller than 20% of the cross-sectional area of the tank, and the effective area of each baffle

should be greater than 40% of the cross-sectional area of the tank". The common shape of baffles is shown in Fig. 1 and the shaded area is baffles. The cross-sectional area in the experiment was 3.14593 m². The area of baffle was 1.451353 m², which was 46.198% of the cross-sectional area of tank, and the area of upper bow was 0.614185 m², which was 19.550% of the cross-sectional area of the tank. They were all in accordance with the rule.

The thickness of the baffle was 0.2 m. So when meshing with FLUENT simulation software, we selected about 0.2 m as the grid size. In order to divide the grid conveniently, the thickness of the baffles was set to be 0.2 m and the volume was 0.290268 m³.

Figures 2, 3, 4 and 5 show the curve of the liquid impact force tank suffered in x-direction with different numbers of baffles when their filling ratios are the same.

We can find that when the filling ratio is $\lambda = 0.25$ m, the number of baffles nearly has no influence on the liquid impact force, because the distance between the bottom edge of baffle and the bottom of the tank is 0.75 m, and the distance between the liquid free surface and the bottom of the tank is $\lambda \cdot \phi = 0.25 \times 0.5$ m. The gap distance between them is 0.25 m, so the baffle failed to block the wave caused by shock then.

The data when the filling ratio is $\lambda = 0.5$, $\lambda = 0.75$ and $\lambda = 0.875$ are shown in Tables 1, 2 and 3.

Above all, increasing the number of baffles can improve the liquid impact effectively, but it will make the peak of impact force become greater. The baffles' effect of defending waves reflects on the impact except for the first peak. Moreover, improving the effect of defending waves is not proportional to increasing the number of baffles. The improvement of effect gets lower with the increase in the number of baffles.

Fig. 1 Shape of the baffles
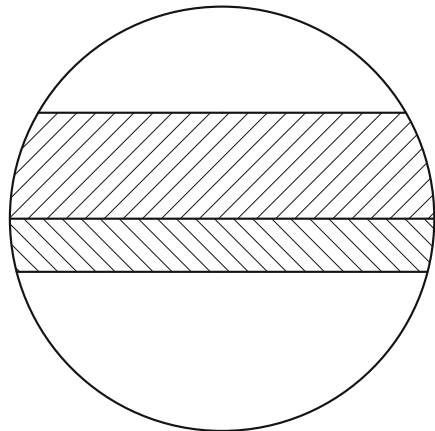
Fig. 2 Curve of the impact force with different numbers of baffles when filling ratio is 0.25
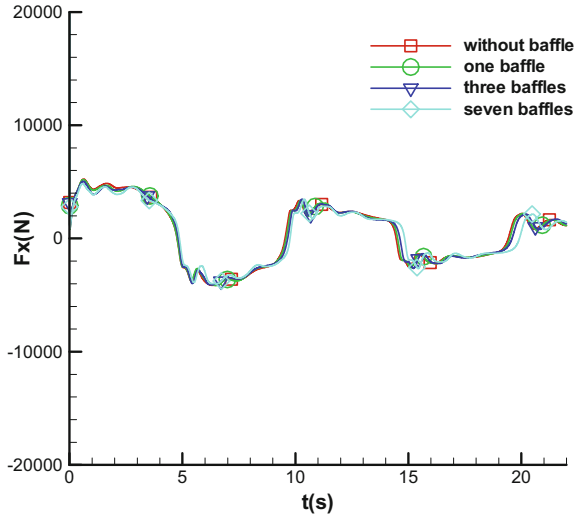
Fig. 3 Curve of the impact force with different numbers of baffles when filling ratio is 0.5
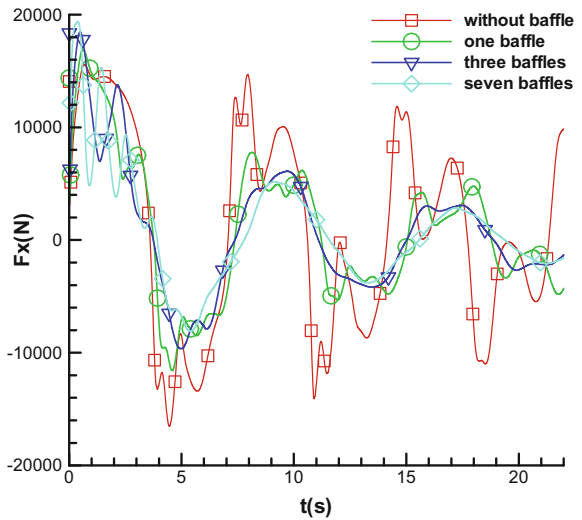
The Influence of Vertical Baffles on Liquid Sloshing

Liquid Simulational Experiment in Tank When the Area of Baffles is 46% of the Cross-sectional Area of Tank

In order to research the influence of effective area of baffles for defending waves (the area below the liquid free surface) on liquid impact force, we set five kinds of

Fig. 4 Curve of the impact force with different numbers of baffles when filling ratio is 0.75
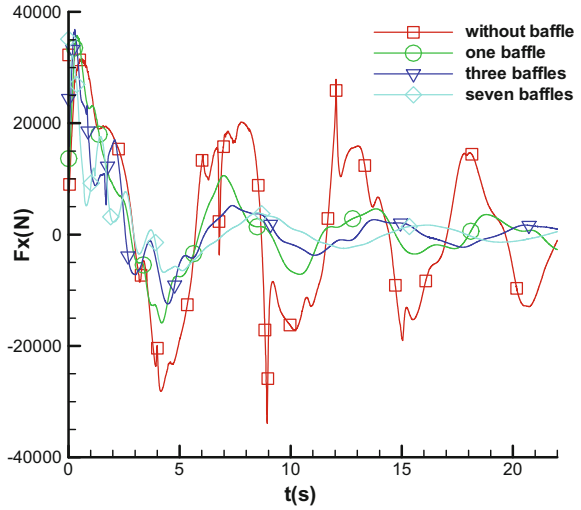
Fig. 5 Curve of the impact force with different numbers of baffles when filling ratio is 0.875
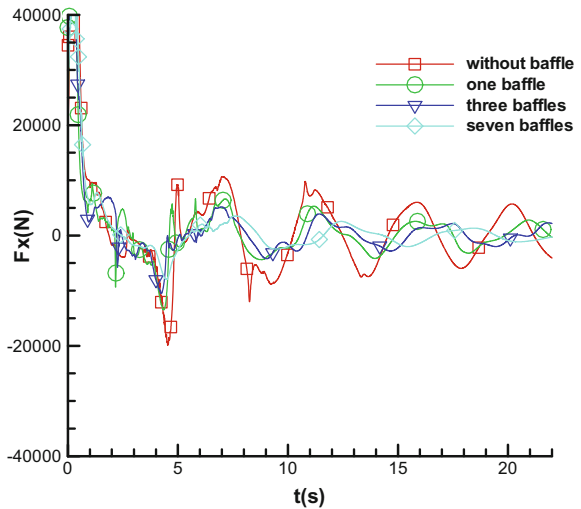
baffles in this section and simulated the influence of them on liquid impact when the filling ratio was 0.75 to find the difference between them. Their location is shown in Figs. 6, 7, 8 and 9. And the area of them was the same, which was 1.451 m².

When the filling ratio was 0.75, the baffles in Figs. 1, 6 and 7 were all under the liquid free surface, in which the area was 1.451 m². The area of baffle below the liquid free surface in Fig. 8 was 1.227771 m², and it was 0.836684 m² in Fig. 9. Figures 10 and 11 show the curve of their liquid impact force.

Table 1 Data with different numbers of baffles when the filling ratio is 0.5

Peak number	The first peak		The first trough		The second peak		The second trough		The third peak	
	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N
No	0.6750	15,569	4.4590	−16,521	7.9500	14,700	10.8900	−14,093	14.5775	11,871
One	0.6625	16,916	4.5900	−11,578	8.1475	7784	11.9450	−5286	17.985	4777
Three	0.4925	18,470	4.9425	−9636	9.6800	6152	13.4400	−4176	17.4475	3082
Seven	0.3850	19,373	5.3825	−8087	9.1175	5172	3.2675	−3818	17.1825	2844

Table 2 Data with different numbers of baffles when the filling ratio is 0.75

Peak number	The first peak		The first trough		The second peak		The second trough		The third peak	
	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N
No	0.6025	31,644	4.155	−21,839	7.7825	20,213	8.9425	−33,897	12.045	27,884
One	0.3825	35,755	4.1925	−15,872	6.9900	10,599	10.4125	−7131	13.8375	4623
Three	0.2825	36,869	4.4775	−12,454	7.3750	5221	11.0800	−3715	13.7575	2677
Seven	0.1775	32,218	4.2900	−6792	8.4925	3916	12.3925	−2489	16.145	1709

Table 3 Data with different numbers of baffles when the filling ratio is 0.875

Peak number	The first peak		The first trough		The second peak		The second trough		The third peak	
	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N
No	0.2025	79,316	4.5375	−19,878	7.0675	10,671	8.2525	−11,979	10.7850	9801
One	0.2300	167,970	4.3575	−13,490	5.8000	6652	8.8125	−4384	11.1850	5317
Three	0.2975	118,733	4.2350	−10,542	6.9800	5193	9.0300	−4210	11.4125	3921
Seven	0.1875	61,016	4.4725	−7962	7.6675	3523	9.4825	−2455	12.445	2563

Fig. 6 Baffles are at the
bottom of the tank
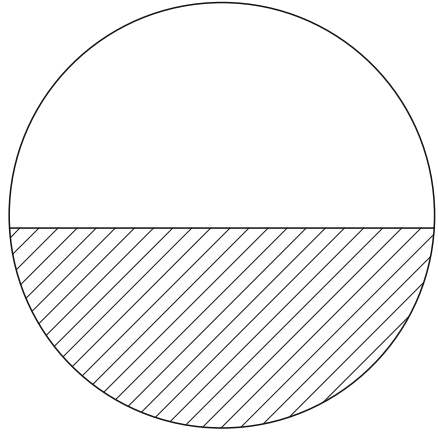
Fig. 7 Baffles are at the low
of the tank
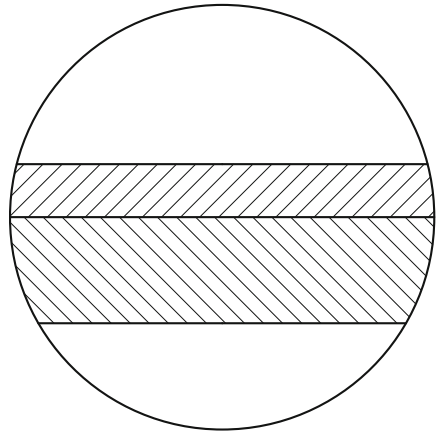
Fig. 8 Baffles are at the high
of the tank
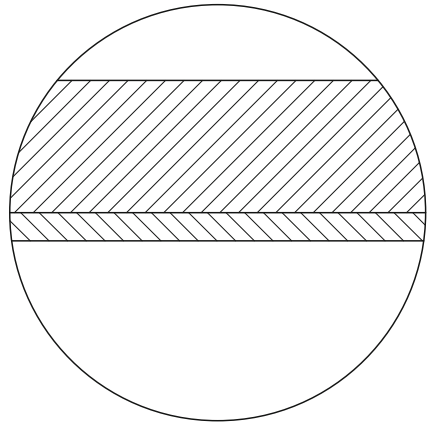
Fig. 9 Baffles are at the top
of the tank
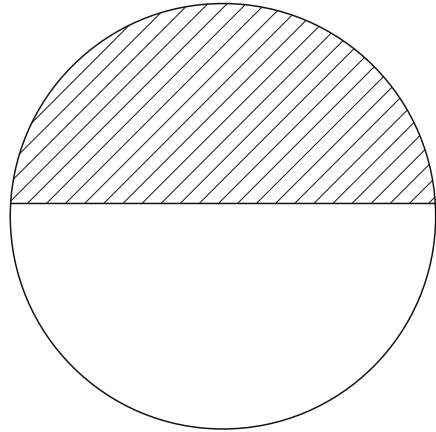
Fig. 10 Curve of five
situations
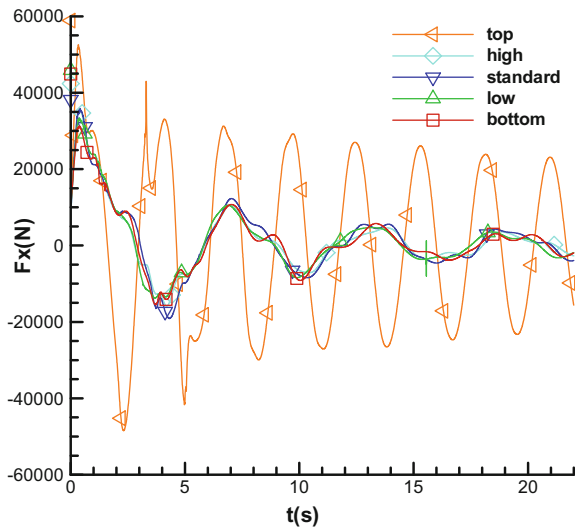
We can find from observing the curve in Fig. 10, the peak of the top one is significantly higher than other curves. And the oscillation frequency is almost double that of others. However, from Fig. 11, we can figure out that instead of defending waves, the top location of baffles exacerbated the liquid impact. Both the peak and the oscillation frequency increase, which lower the vehicle security.

The waves of other four curves are almost unanimous, and it is difficult to distinguish between their advantages and disadvantages.

Fig. 11 Curve of the top and
without a baffle
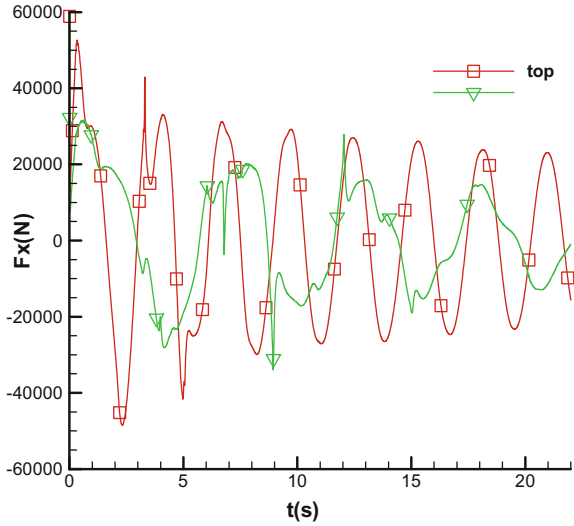
Liquid Simulational Experiment in Tank When the Area of Baffles is 20% of the Cross-sectional Area of Tank

We set four kinds of baffles in this section and simulated the influence of them on liquid impact when the filling ratio was 0.5 to find the difference between them. The area of four kinds of baffles was the same, which was 0.614 m². And their location is shown in Figs. 12, 13, 14 and 15.

The curve of several cases is shown in Fig. 16. The curve of top one is essentially coincident with the curve of the no baffles one. Because the top location of baffles couldn't play a role in defending waves when the filling ratio is 0.5.

Fig. 12 Baffles are at the
bottom of the tank
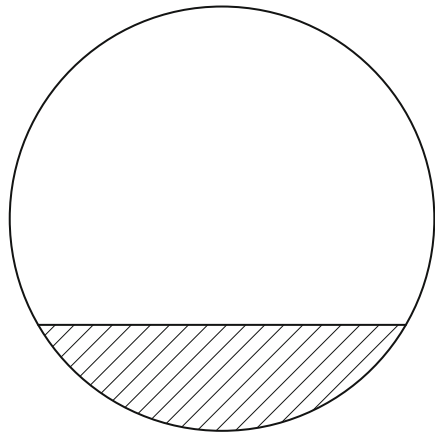
Fig. 13 Baffles are at the low
of the tank
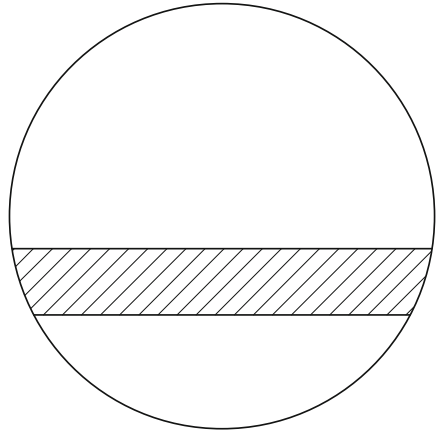
Fig. 14 Baffles are at the
high of the tank
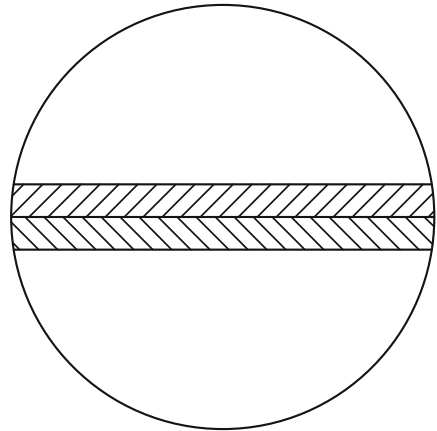
Fig. 15 Baffles are at the top
of the tank
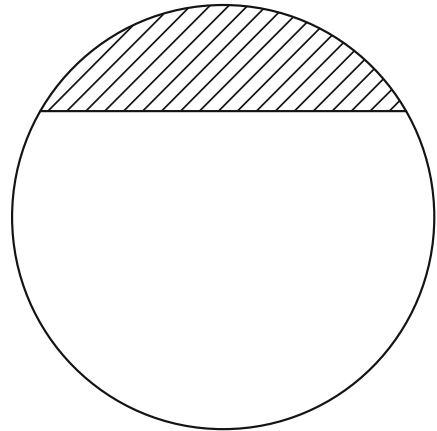
Fig. 16 Curve of four
situations
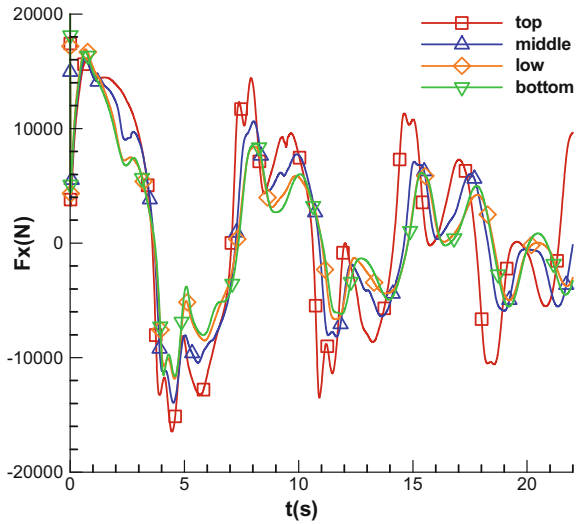
Fig. 17 Curve of the top and
without a baffle
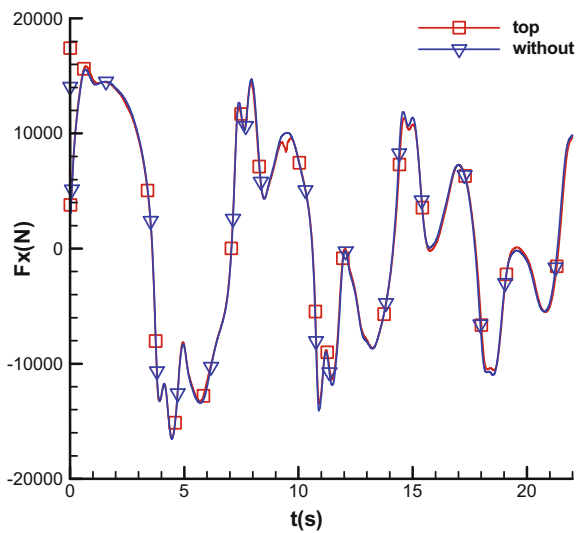
However, the area of the low and the bottom locations is both totally under the free
surface of the liquid, so both curves are essentially coincident. Since the shape of
the middle one was similar to the low one, the area below the liquid free surface
was less than the low one by 50% and the effect of defending waves also was not as
good as the former. So we can point out that the area below the liquid free surface is
the most important factor of defending waves (Fig. 17).

Three Baffles are Interlaced

Figure 18 illustrates three baffles staggered by 60°, and they are, respectively, arranged on three-quartered cross section in x-direction. The width of the baffles was 0.743 m, the baffles' area was 1.451074 m^2, the thickness was 0.2 m and the volume was 0.290216 m^3. The area, thickness and the volume were consistent with the standard one. When the filling ratio was $\lambda = 0.5$, the area below the liquid free surface was 2.176611 m^2, which was 50% of the total area, while the total area below the liquid free surface of three standard-set baffles was $0.494742 \times 3 = 1.484226$ m^2. The former area was higher than the latter by 46.69%.

The curve of the liquid impact force of the interlaced and the standard arrangement of baffles is shown in Figs. 19 and 20.

Fig. 18 Baffles are interlaced with 60°
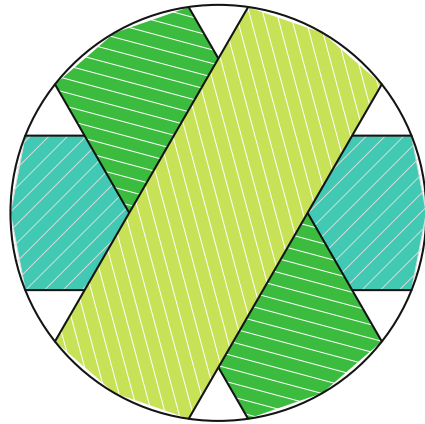
Fig. 19 Curve of the interlaced and the standard setting of baffles when the filling ratio is 0.5
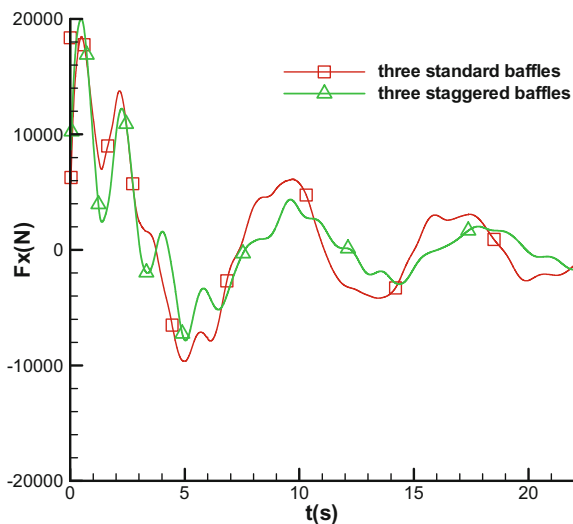
Fig. 20 Curve of the interlaced and the standard setting of baffles when the filling ratio is 0.75
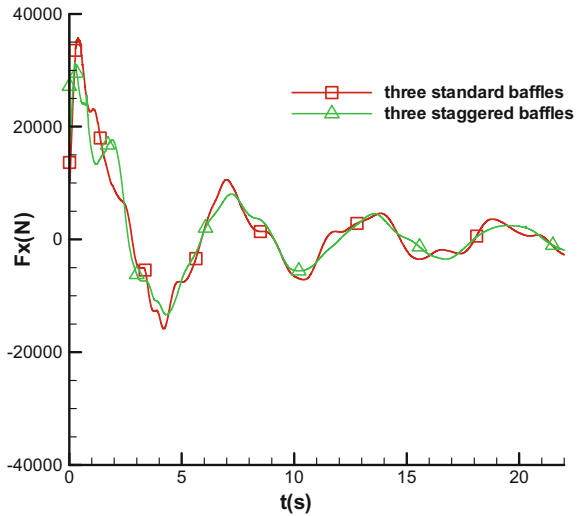
Table 4 Data of interlaced baffles when the filling ratio is 0.5

$\lambda = 0.5$	The first peak		The first trough		The second peak		The second trough		The third peak	
	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N
Standard	0.4925	18,470	4.9425	−9636	9.6800	6152	13.4400	−4176	17.4475	3082
Interlaced	0.4775	20,026	5.0350	−7837	9.6125	4363	14.3625	−2963	17.8075	2031
Increasing		8.24		−18.67		−29.08		−29.05		−34.10

Table 5 Data of interlaced baffles when the filling ratio is 0.75

$\lambda = 0.75$	The first peak		The first trough		The second peak		The second trough		The third peak	
	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N	Moment/s	Peak/N
Standard	0.2825	36,869	4.4775	−12,454	7.3750	5221	11.0800	−3715	13.7575	2677
Interlaced	0.2625	31,267	4.3250	−13,361	7.2150	8051	10.1900	−5565	13.5925	4558
Increasing		−15.19		7.28		54.20		49.80		70.27

The data of interlaced baffles when the filling ratio is 0.5 and 0.75 are shown in Tables 4 and 5.

To sum up, baffles interlaced with 60° act well in defending waves when the filling ratio is 0.5, while the standard one acts better when the filling ratio is 0.75. Comparing two types of area below the liquid free surface, we can find that the most important factor which influences the liquid impact is the area used to intercept liquid impact.

Conclusions

a. In circle cross-sectional tanks, increasing the number of baffles can efficiently reduce the liquid longitudinal impact. But as the number of baffles increases, the speed of the effect of defending waves begins to slow down. The effect of defending waves is not proportional to increasing the number of baffles.

b. According to the simulational results of different locations of baffles, we come to the conclusion that the area below the liquid free surface is the most important factor which influences the effect of defending waves. The bigger the area below the liquid free surface, the better the effect of defending waves.

References

1. Wan, Liping, and Caifu Qian. 2012. Design of swash plate for liquefied-gas tank. *Chemical Machinery* 01: 48–51.
2. Gao, Bingjun, and Chengwen Shang. 2012. Impant of different location of baffles on baffling effect during the tank braking. *Journal of Hebei University of Technology* 03: 46–49.
3. Zeng Yue, Nan Jiang, Wubin Chen. 2013. Numerical simulation study on anti-wave effect of baffle. *Process Equipment & Piping* 05: 18–23.
4. Chen, Zhiwei. 2006. Numerical Simulation of Liquid Sloshing in Transportable Pressure Vessel and Research on the Baffles.
5. Cacciatore, Philip J., Benjamin F. Hantz Iv, and L. Magnus Gustafsson. 2008. Evaluation of storage tank floating roofs for stress and stability due to earthquake induced liquid sloshing. San Antonio, TX, United States: American Society of Mechanical Engineers.
6. Cheng, Xu-Duo, Yong Liang, and Ji-Hua Wen. 2004. Hydroelastic vibrations and liquid sloshing suppression in a rectangular tank with elastic spacer. *Journal of Hydrodynamics* 16 (3): 336–340.
7. Firouz-Abadi, R.D., H. Haddadpour, and M. Ghasemi. 2009. Reduced order modeling of liquid sloshing in 3D tanks using boundary element method. *Engineering Analysis with Boundary Elements* 33 (6): 750–761.
8. Gradinscak, Marija, and Farial Jafar. 2013. *Computational modelling of liquid sloshing in rectangular tank: Jeju Island*. Republic of Korea: Trans Tech Publications Ltd.
9. Hou, Ling, Fangcheng Li, and Wu Chunliang. 2012. A numerical study of liquid sloshing in a two-dimensional tank under external excitations. *Journal of Marine Science and Technology* (Japan) 17 (3): 305–310.

Research on the Collaborative Management and Information Service System of Comprehensive Passenger Transportation Hub

Weiwei Li and Yuguang Sun

Abstract Comprehensive passenger transport hub (CPTH) is the point of different transport modes of the transportation network. Information integration is one of the basic features of modern CPTH. Directing at the problems of passenger fast transfer guide, large passenger flow monitoring, and early warning and emergency evacuation of CPTH, the paper proposes a collaborative management and information service system of CPTH. Based on the analysis of demand, all the essential elements of system construction are analyzed, and the system functional architecture of the system is designed. Then, the logical relationship structure of all the application systems is proposed. At the same time, this paper introduces the main function design ideas and main points of all application subsystems. Finally, the realization of each software function module is described in detail. System application results show that the system can effectively improve the travel efficiency for passengers and promote the development of comprehensive transportation system.

Keywords CPTH · Information system · Collaborative management · Information service

Introduction

CPTH is a collection of long-distance passenger transport, the city public transport (often including rail, bus, and taxi.), and private transportation. Integration of infrastructure, information construction, and operation mechanism are the basic features of modern comprehensive passenger transport hub (CPTH). At present, Beijing, Shanghai, Guangzhou, Shenzhen, and other major cities in China have been built in a number of large-scale urban integrated transport hubs. In order to maximize overall operational efficiency of CPTH and improve the travel efficiency

W. Li (✉) · Y. Sun
Research Institute of Highway Ministry of Transport, Beijing 100088, China
e-mail: weiwei.li@rioh.cn

for passengers, research of the information system for CPTH has attracted much more research attention [1, 2].

In recent years, several researches on information system for CPTH have been done. Ling and Ye [3] present the travel information sharing platform concept, which integrates the related MIS of various traffic modes effectively, and realize the exchange and sharing of information. Lin et al. [4] present a hub information system framework, which is based on the information exchange platform of the hub, the application system environment for the integrated traffic information service of the passenger and the integrated management and collaborative work of the hub is formed. Yang [5] highlights the core role of operation management system in the integrated transportation hub operation and management platform. Based on the spatial data, library data, and integrated transportation information database, the operational status of the integrated transportation hub is displayed in real time through the GIS shared service platform and the application service platform to integrate the business information of the integrated transportation hub.

CPTH as a variety of modes of transportation integration, the convergence of the organic whole, in the construction of the information system must be comprehensive consideration of various traffic modes of transfer and join, induced traffic, traffic organization, public areas of collaborative management and other factors, the various modes of transportation operation coordination linkage, optimize the allocation of resources, traffic demand balance, and unified emergency command. This study proposes the collaborative management and information service system of CPTH, through the establishment of covering a variety of modes of transport services and management information sharing and exchange platform, to realize collaborative operation in comprehensive passenger transport hub long-distance passenger transport, city bus, taxi, rail transportation and other different transport modes, and implementation of integrated passenger transport hub in case of unexpected emergency command and coordination; provide internal and external integrated passenger transportation terminal transfer, import stations, and other traffic information service; promote the organic, efficient convergence between the various modes of transport, improving passenger transfer efficiency; reduce passengers delays; enhance the ability to deal with the construction of emergency situations; and improve traffic monitoring and early warning, efficient organization, and emergency evacuation efficiency.

System Framework Design

The collaborative management and information service system of CPTH should be to strengthen connection mode of transport capacity, reduce the public transfer time and space distance, and promote different transport modes, operation and management of the cohesion, and coordination efficiency promotion as the basic principle, using modern information technology to achieve in the hub multitransportation mode of collaborative management and integrated services.

System should have intelligent depot management, station induction and service information, monitoring the daily operation of hub, comanagement and joint support, safety evacuation and emergency management, and comprehensive running information management function and meet the comprehensive transfer hub passenger station daily operation and management, safety in production, and the government sector management needs.

The collaborative management and information service system's functional architecture of CPTH is shown in Fig. 1. From bottom up, the six layers are physical place layer, infrastructure support environment layer, data resources layer, application system layer, application service presentation layer, and user layer. The left and right sides are the security system layer, includes standard specification system and information security guarantee system.
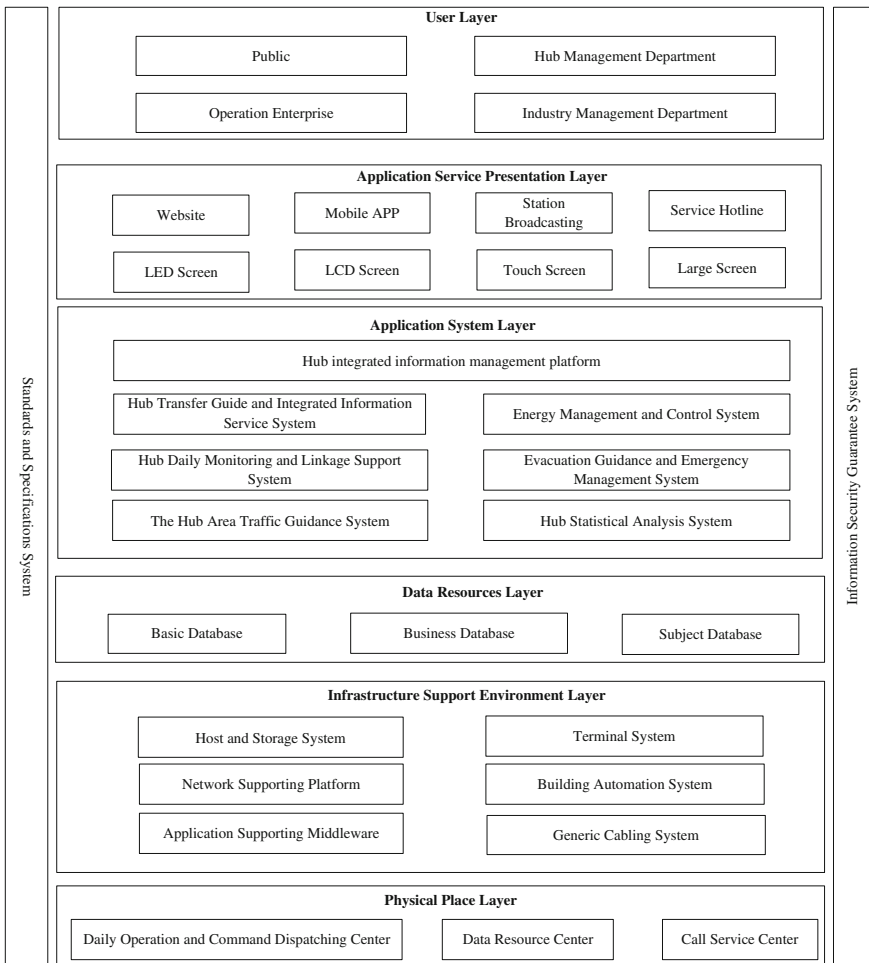
Fig. 1 System functional architecture for CPTH

Physical Place Layer

The construction of the physical site includes daily operation and command dispatching center, data resource center decoration, and mechanical and electrical engineering. Daily operation and dispatch center is to provide physical place for passenger station daily operation scheduling and emergency command. The data resource center is to provide physical places for application system deployment.

Infrastructure Support Environment Layer

Infrastructure support environment layer provides the software and hardware infrastructure supporting environment for the data resources and application system. The Infrastructure support environment layer mainly comprises host and storage system, application supporting middleware, network supporting platform, terminal system, building automation system, and generic cabling system.

Data Resources Layer

The data resources layer mainly includes the basic database, the business database, and the subject database. Basic database mainly includes the hub infrastructure data, electromechanical equipment's data, and emergency resource data. Business database mainly includes hub daily running data, dynamic supervision inspection database, and emergency events data. Subject database mainly includes daily feature theme database, holiday feature theme, and emergency feature theme database.

Application System Layer

This layer is composed of a set of application software, which includes one platform and six systems. All the application software shall be developed based on uniform software architecture and a uniform information coding system. The system should adopt the service-oriented architecture (SOA), reasonably divide the function module's granularity, and adopt the ESB bus mode to realize the interconnection and flexible access of each function module.

Application Service Presentation Layer

Application service presentation layer consists of Web site, mobile APP, LED or LCD display terminal, broadcast service, touch screen, telephone service, and other

means achieves a full range, multiway, humanization, a station type passenger station transfer guiding, and provides the integrated information service inside and outside the hub.

User Layer

User layer consists of social public, operation enterprises (such as railway, airport, subway, long-distance passenger transportation, and taxi companies), hub management department, transportation industry management department, public security management, and urban management departments.

Security System Layer

Security system layer includes standard specification system and information security guarantee system. Standard and specification system lays the foundation for the future expansion and application. Information security guarantee system provides security support for the above five basic layers.

Application System Design

The application system emphasizes to provide a safe, convenient, and efficient travel service for passengers as the main line and is committed to the integrated application of each subsystem.

The logical relationship diagram of application system is shown in Fig. 2.

Hub Transfer Guide and Integrated Information Service System

Hub transfer guide and integrated information service system is based on 5R as the principle, seamless integrated transfer as the goal. The 5R principle is to provide the right information services for the appropriate population on the right time and at the right place in an appropriate manner. The system is mainly oriented to the demand for transfer in the hub station [6]. At the same time, the scope and means of information services are extended to provide a full range of travel information services for the travelers according to the different travel periods. Pre-trip information services mainly through the portal Web site, mobile APP, or service telephone to provide pre-trip planning services. Travel information service mainly uses
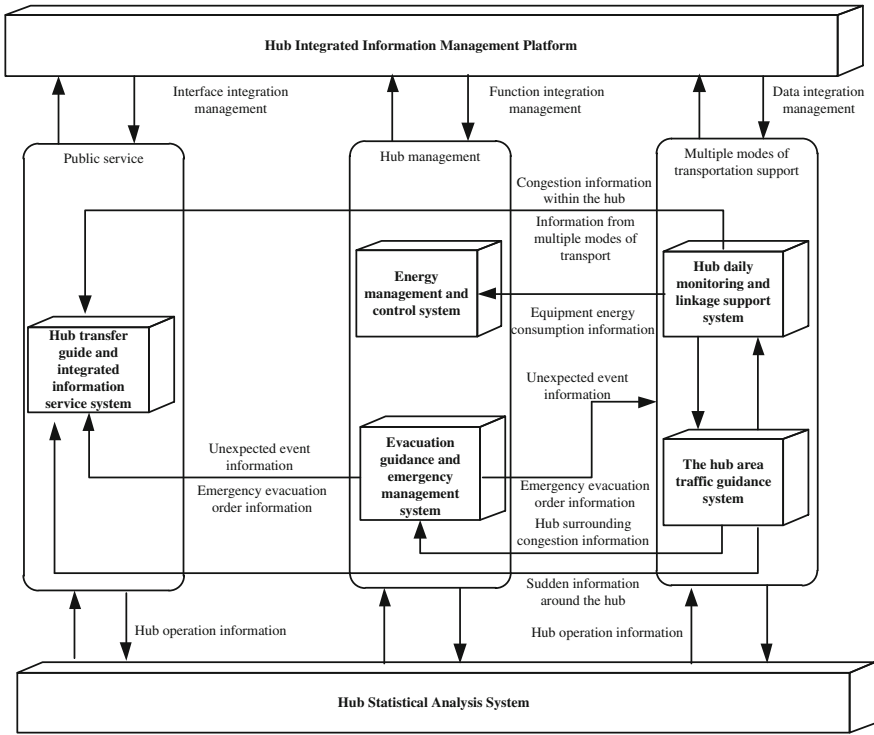
Fig. 2 Logical relationship diagram of application system

static and dynamic display terminal to provide transfer, delays, and other information. After the completion of trip, travelers can get lost inquiries or complaints information service by service hotline or Web site.

Hub Daily Monitoring and Linkage Support System

Hub daily monitoring and linkage support is mainly for hub management needs and realize the transfer hub area, office area, public bus hub station, taxi station area, public area, and all-round monitoring of the refueling station, maintenance work area, perimeter, and other physical hub operation area. Through comprehensive analysis and forecast data on monitoring facilities, operation safety monitoring data and environmental monitoring information transfer the various modes of transport passenger flow and a variety of modes of transport to meet the needs of the operation and management of transportation hub. At the same time, implementation of the information exchange and sharing between other modes of transportation to achieve a variety of modes of transport traffic hub emergency coordination linkage and to improve the operation management level of traffic hub.

The Hub Area Traffic Guidance System

The hub area traffic guidance system is mainly oriented to the information service demand of the vehicle drivers for import and export hub, including the parking guidance function and the outbound route guidance function. By setting the information collection equipment in the underground parking lot and the taxi ride down area, the information of the parking lot and the information of parking spaces in the parking lot are obtained, which will be released to the public vehicles or taxi drivers entering the hub. In addition, the system will also get hub surrounding the real-time traffic information from the urban dynamic traffic information collection system and guide the vehicle driver out of the station rational planning travel routes through the information release equipment is arranged at the outlet of the parking field, which can greatly ease the traffic congestion around the hub station [7].

Energy Management and Control System

Energy management and control system is the hub station buildings within the substation and power distribution, lighting, elevators, air-conditioning, heating, drainage, and other energy usage, centralized monitoring management, and decentralized control system. The system is a general designation of hardware system and software system for building energy consumption on line monitoring and dynamic analysis function. It consists of various measurement devices, data acquisition, and energy consumption data management software system.

Evacuation Guidance and Emergency Management System

Evacuation guidance and emergency management system is oriented to the needs of hub emergency management. System is to achieve the rapid alarm, quick response, and scientific decision-making function of public emergency. The system mainly realizes the centralized control of the safe evacuation information, so as to quickly establish the safe evacuation channel, realize the rapid evacuation of passengers and staff in the hub, and ensure the safety of the public.

Hub Statistical Analysis System

Hub statistical analysis system mainly obtains the key operation and the passenger flow rule and realizes statistics and analysis on the comprehensive operation data. The system generates all kinds of statistical information detailed reports, to provide data support for the operation and management of the hub.

Hub Integrated Information Management Platform

Hub integrated information management platform is the centralized management platform of six application systems, which can realize the user interface integration, data integration, and function integration of the above six application subsystems.

System Function Module Design

Hub Transfer Guide and Integrated Information Service System

(1) *Information organization and management*: The information that needs to be released is organized in a manner that is more than one dimension and is organized in a hierarchical manner.
(2) *Release terminal management*: For the management of distribution terminal equipment type, name, type, location, equipment status, display mode (text or graphics or image), and text size specifications.
(3) *Information collection and management*: Relying on information collection terminals or manual submission and other ways to achieve the collection of all kinds of release information.
(4) *Information processing and editing*: Used for the collection of all kinds of information for the deep application of processing under what you see is what you get(WYSIWYG) mode.
(5) *Information audit management*: Used for the collection of information automatically prompted and manual audit confirmation.
(6) *Information release management*: Used in accordance with the provisions of the information organization and management, and the relevant information is sent to the information release terminal.

Hub Daily Monitoring and Linkage Support System

(1) *Daily monitoring management*

   (a) *Equipment operation condition monitoring*: For the hub building environment parameters and ventilation and air-conditioning, electric power, lighting, to drainage, elevators and other mechanical and electrical equipment or system operation state, the terminal equipment (such as LCD and LED system), running state monitoring, and early warning.

(b) *Video monitoring*: Used to realize the key areas of transport hub and transfer area security monitoring. Transportation hub management, industry management, and traffic law enforcement agencies should have the right to control the video surveillance system, the right to configure, in order to facilitate remote monitoring, video correction violation, and emergency response.

(c) *Passenger flow monitoring*: For the realization of the main entrances and exits, transfer channels and other key areas of the passenger flow monitoring and early warning.

(d) *Passenger behavior analysis*: Used to realize the traffic hub in the passenger flow drama or sensitive areas of the density of passenger flow, movement characteristics, and other acts of detection and early warning.

(2) *Coordinated decision-making support*

(a) *Daily coordination scheduling*: Sharing information exchange in various modes of transportation hub, to provide decision support for the coordination of various modes of transport.

(b) *Emergency linkage control*: Realize the coordinated transportation hub in the emergency incident.

The Hub Area Traffic Guidance System

(1) *Parking guidance function*

(a) *Position guide*: Used to indicate the hub location and parking lot entrance direction, which is arranged in the parking field at the entrance of the hub static logo implementation.

(b) *The calculation of free parking lot*: Through the acquisition of the main entrance of the parking lot hub into and out of the vehicle data, calculate the hub parking lot's free parking spaces and additional data, location data.

(c) *Free parking lot parking space allocation*: The system through the identification of vehicles position, combined with the distribution of free parking, automatic extraction of vehicle distance into the nearest parking number and prompt the driver parking locations roughly.

(d) *Release free parking information*: The system automatically calculate the free parking data, issued to social vehicles driving around the hub using LED display terminal equipment and to achieve dynamic parking information service function.

(e) *The surrounding traffic information release*: The system mainly serves the social vehicles leaving hub parking lot and driving around the hub. The system displays information through LED display terminal set nearby the parking lot exit through the real-time traffic information acquisition.

(f) *Query and statistics*: Based on historical data, according to the different dimensions of keywords and time to generate statistical reports, calculate the different dimensions of the parking lot utilization rate.

(2) *Taxi guidance function*

(a) *Position guide*: Used to indicate the location and direction of the taxi ride and drop zone in the hub, by setting the static identification of the entrance of the taxi ride and drop zone.

(b) *License plate recognition and taxi number statistics*: For license plate recognition and counting function to stop and taxi station. At the same time, including the identification of the blacklist of the vehicle entering, leaving the taxi number statistics.

(c) *The number of passengers waiting statistics*: By setting the area by lowering the waiting area at the top of the camera in the taxi, the taxi waiting area of the video image acquisition and analysis software for video image processing and analysis using video intelligence and the number of taxi waiting area statistics.

(d) *Taxi spare parking calculation*: Through the acquisition of a taxi by lower area of entering and leaving the taxi data, counting the taxi passenger drop zone spare the number of parking spaces and real-time transmission to the information release module.

(e) *Taxi-free parking information release*: This module implements dynamic taxi parking information service function.

Energy Management and Control System

(1) *Daily energy consumption monitoring and management function*

(a) *Monitoring and management of electricity usage*: Monitoring and management of electricity consumption of each monitoring point.

(b) *Monitoring and management of water use*: Monitoring and management of water consumption for each monitoring point.

(c) *Heating monitoring and management*: Heating monitoring and management of each monitoring point.

(2) *Energy consumption statistics analysis and early warning function*

(a) *Energy consumption analysis*: Analysis of energy consumption for each monitoring point and transportation hub.

(b) *Early warning of abnormal energy consumption*: By monitoring and analyzing the energy consumption of each monitoring point, the abnormal situation of energy consumption in time.

(c) *Information inquiry*: For each monitoring point and transportation energy consumption monitoring information query.

(d) *Energy consumption statistics*: Statistics for each monitoring point and the energy consumption of the transport hub.

(e) *Information publicity*: Used to display each monitoring point data of the traffic hub.

Evacuation Guidance and Emergency Management System

(1) *Emergency resource management*: Including emergency agencies management, experts and emergency facilities or equipments management, emergency files management, emergency knowledge management, and the video monitoring resource management.

(2) *Plan, exercise, and training management*: Including the management of various types of emergency response plan, exercise information management, exercise evaluation management, training information management, training evaluation, and management functions.

(3) *Daily duty and quick submission*: For daily management, personnel will be timely reporting to the relevant departments in a timely manner. Including mail list management, duty management, message management, quick alarm, and rapid reporting function.

(4) *Emergency response decision-making*: Including event assessment, start planning or development program, on-site video surveillance, on-site command, event relief, and other functions.

(5) *Centralized control of safe evacuation*: Used to achieve a unified and centralized control of the relevant terminal equipment, in order to facilitate the safety and rapid evacuation of passengers and staff.

(6) *Post disposal management*: Including the approved emergency caused by the loss, all kinds of personnel and equipment investment, pension, and compensation scheme.

Hub Statistical Analysis System

(1) *Data analysis of the daily hub passenger flow*: Including the hub passenger boarding and alighting statistics of different periods and pear hour, statistical analysis of typical period hub passenger flow characteristics such as on weekends and holidays and other statistical analysis of passenger characteristics of abnormal conditions: bad weather, emergencies (such as passenger rail transportation hub station outage), public health events and the surrounding road construction period analysis of real-time traffic data hub with the historical period of data comparison, passenger traffic trend estimation, and abnormal warning .

(2) *Analysis of operation data of each mode of transport statistics*: In the hub of
passenger flow by means of transport (bus, long-distance passenger transit, and
rail transit) to carry on the statistical analysis of operation data. Including the
class of different amount of periodic line statistics, analysis on the peak hour
class of statistics, and statistical analysis on weekends and holidays, the typical
period shift class line dispatching law, abnormal conditions to shift scheduling
and statistical analysis, operation data compared with the same period of his-
torical data analysis.

(3) *Analysis of safety operation of hub*: Including the statistical analysis of emer-
gency data hub according to different time periods or events categories, and the
hub pit operating vehicle violation records the violation rate statistics. These
statistics include the early warning, event information query, according to the
time, event type, event level conditions such as the combination of the number
of events, property damage, and compensation cost indicators for statistical
assistance.

Hub Integrated Information Management Platform

(1) *Integrated user interface management*: Using the unified user management,
unified authentication, and single sign on technology, a unified integration of
subsystems realizes the unity of the performance level of system.

(2) *Integrated data management*: To achieve all kinds of information resources in
the hub of fully integrated and unified management, as well as the related traffic
information resource integration access. Based on the data exchange platform to
realize data exchange time and fixed batch data sharing. Real-time data service
is based on the real-time message middleware.

(3) *Integrated function management*: Based on the design principle of tightly
coupled data, loosely coupled business, to achieve business linkage system.

Conclusions and Recommendations

The establishment of collaborative management and information service system of
CPTH is to accelerate the application of information technology integration and
innovation in each hub construction and management and other aspects. The pro-
duction of this paper has been applied into information construction of CPTH of
some cities of China. System practical application results show that this system can
effectively improve the travel efficiency for passengers and promote the develop-
ment of comprehensive transportation system. Meanwhile, the system construction
is a large system engineering, which needs to integrate all kinds of information
systems, gathering all kinds of information resources. The further research should

focus on how to realize the more efficient and accurate collection and prediction of passenger flow data, and how to realize the sharing and exchange of information resources, timely update, dynamic publishing, and efficient query.

References

1. Zhang, Xiaozheng, Yi Zhang, and Caijun Du. 2013. Information service and dynamic induce technology used in urban integrate passenger transport hub. *China ITS Journal* 1: 126–128.
2. Cao, Jiandong, Zhenhua Liu, and Yingjie Ma. 2015. Promote information technology development of integrated transport system in China. *China Transportation Review* 37 (8): 38–45.
3. Ling, Qiao, and Qing Ye. 2011. Travel information sharing platform for comprehensive passenger hub. In *Third International Conference on Transportation Engineering (ICTE)*, 781–786.
4. Lin, Guoxin, Shigui Luo, and Cong Miao. 2012. Requirement analysis and framework research on information system for comprehensive passenger transport hub. *Highway* 5: 239–243.
5. Yang, Fang. 2015. Exploration of comprehensive transportation hub management platform construction. *Building Technique Development* 42 (10): 64–66.
6. Jiang, Lijuan, and Minhua Shao. 2014. Research on traveler information need in the interior of large-scale multi-modal transport hub. In *China 14th COTA International Conference of Transportation Professionals*, 1327–1337.
7. Chen, Min, and Xuesong Zhang. 2009. The traffic guidance system of elevated roads in Hongqiao comprehensive transportation hub. In *Second International Conference on Transportation Engineering*, 2401–2406.

Decision-Making Model of Lane-Change Behavior Based on Integrated Cognitive Vehicle Cluster Situations

Jinglei Zhang, Xiaoyuan Wang, Jianqiang Wang and Jingheng Wang

Abstract Lane-change behavior is the most important factor that explains changes in vehicle cluster situations, and it is of great importance to traffic safety. It plays a significant role in research pertaining to traffic flow and theoretical studies about active vehicle safety. This is particularly prevalent in intelligent (auto- and assisted-) driving systems, where decision-making processes with respect to lane-change behavior can be identified. Based on previous studies and by considering the impact of vehicle cluster situations on task prioritization with regard to comprehensive cognitive reactions of lane-change processes, a new lane-changing decision-making model was developed. This model considers a range of factors, including the propensity of the driver of the target vehicle and the characteristics of surrounding vehicles (e.g., decision-making, distance, and speed). This model supports the recreation of the complex system wherein vehicle clusters are found. The field data, virtual reality data, and simulation data were used to validate the aforementioned model. It was found that the modeled results can provide guidance on whether drivers will decide to change lanes.

J. Zhang · X. Wang (✉)
School of Transportation and Vehicle Engineering, Shandong University of Technology, Zibo 255091, China
e-mail: wangxiaoyuan@sdut.edu.cn

X. Wang · J. Wang
State Key Laboratory of Automotive Safety and Energy, Tsinghua University, Beijing 100084, China

J. Wang
Shandong Zibo Experimental High School, Zibo 255000, China

Introduction

Car following and lane change are the two most important driving tasks while on the road and are essential components of traffic flow and theoretical research on active vehicle safety. Car following and lane change describe the longitudinal and lateral interactions of vehicles on the road, respectively. Compared with car-following behavior, more factors need to be considered in the study of lane-change behavior, and the decision-making processes of drivers are considered to be more complex.

Traditional lane-changing models mainly simulate the associated decision-making process based on mathematics and logical rules. However, drivers' decisions are always based on approximate perceptions of the traffic around them. With respect to this perspective, Wu Jianping et al. [1, 2] presented a traffic flow simulation model of UK motorways based on fuzzy logic. All instances of lane changes were classified as lane changing either from the nearside lane to the offside lane (LCO) or from the offside lane to the nearside lane (LCN) in the model. The purpose of LCO is to have some form of speed-based advantage, while LCN is to reduce the impedance to fast-moving vehicles approaching from the rear. The "overtaking benefit" and "lane-change opportunity" are two input variables of the LCO rule base. The overtaking benefit is measured as a function of speed gain and the associated safety and comfort of changing lanes as measured by the lead time to the nearest approaching vehicle to the rear in the offside lane. The "pressure from rear" and "gap satisfaction" in the nearside lane are two input variables of the LCN rule base. The variable "pressure from rear" is measured by the lead time of the following vehicle(s), and "gap satisfaction" is measured by the driving time of target vehicle without reducing the speed. Das and Bowles [3] proposed a new microscopic simulation methodology that divides lane-changing maneuvers into mandatory lane changes (MLCs) and discretionary lane changes (DLCs) based on fuzzy IF–THEN rules. MLC occurs when a driver is forced to leave the current lane, for instance, when merging onto the freeway from an on-ramp or taking an exit off-ramp. DLC is performed when the driver is not satisfied with the driving situation in the current lane and wishes to gain some speed advantage, for instance, when the driver is obstructed by a slow-moving vehicle. The distance of a target car and the ramp metering and the numbers of lanes that need to be crossed are defined as input variables for MLC fuzzy rules. The more lanes that need to be crossed, the higher probability of lane changes will occur. The degree of satisfaction with the current speed and overcrowding of adjacent lanes are defined as input variables for DLC fuzzy rules. Errampalli et al. [4] defined speed advantage, feasibility level, remaining distance to target turn, distance to leader vehicle, and distance to follower vehicle as input variables and the intention of lane change as an output variable to describe the fuzzy decision-making process of lane changing when in speed advantage purpose, turn at intersection purpose, priority bus lane policy purpose, and bus stop/parked vehicle purpose, respectively. Moridpouri et al. [5] present a fuzzy logic lane-changing decision model for heavy-vehicle drivers on freeways

and divided the motivation of lane changing into right changing and left changing. The explanatory variables of right changing include the front space gap, the rear space gap, the lag space gap in the right lane, and the average speed of the surrounding vehicles in the current lane. The explanatory variables of left changing include the front relative speed, the lag relative speed in the left lane, and the average speeds of the surrounding vehicles in the current and the left lane. The trajectory data of heavy traffic conditions were used to verify the model, and the results show that the new fuzzy logic lane-changing decision model could closely replicate the observed lane-changing decisions of heavy-vehicle drivers.

Some researchers used the fuzzy logic approach to study MLC, DLC, and left or right lane changing. These models only considered the impact of front, rear and side lateral front, rear vehicles on target vehicle lane-changing decisions but did not consider the impact of the target vehicle and surrounding vehicle cluster situation to task concentration comprehensive cognitive reaction activity on the process of lane change. In light of this gap, a comprehensive, lane-changing decision-making model that considers the target vehicle drivers' propensity and the surrounding vehicles' types, distance, and speed in complex vehicle cluster environments was developed. It is considered to be important for the study of traffic flow and the theory-based investigation of active vehicle safety.

Analysis of Vehicle Cluster Situation and Lane Changing

Vehicle Cluster Situation

Traffic situation refers to the states and situations of all traffic units that are deployed or in motion and the behavior in the drivers' interest-sensitive area and it contains all of the information about the traffic entities [6]. Here, an interest-sensitive area is defined as an area that has a greater impact on vehicle safety and that requires more of the drivers' attention to the road. The traffic situation includes the characteristics of traffic entities, vehicle group relations (that arise when drivers aim to complete their driving tasks), and scene interpretation information. The characteristics of traffic entity are related to the vehicle types, velocity, acceleration, and the types of drivers. Vehicle group relation describes the vehicles' changing positional relation, which is organized through clustering behavior and the evolution of self-organization. Scene interpretation information describes the natural, artificial, and social environment where a car is located. In this study, the vehicle cluster situation is used as an example (including the characteristics of traffic entity and vehicle group relation) wherein a driver's lane-changing behavior in an interest-sensitive area can be investigated.

The changes in vehicle cluster situations influence the drivers' driving behavior and vice versa. The most important reason that results in changes in vehicle cluster situations has been attributed to lane change. Under different vehicle cluster

situations, in particular with different road usage, different number of lanes, different lanes, which the target vehicle is driving on, and different target vehicle drivers' propensity, have variable impacts on the lane-changing decision-making processes of the target vehicle driver.

Lane Changing

Take a three-lane road as an example. When the target vehicle is located in the middle lane, according to the position of the target vehicle's front bumper, the interest-sensitive area is divided into left front, left rear, front, rear, right front, and right rear subregions.

In Fig. 1, the speed of target vehicle n_1, left front vehicle n_2, left rear vehicle n_3, front vehicle n_4, rear vehicle n_5, and right front vehicle n_6 are denoted as v_{n_i} ($i = 1, 2, 3, 4, 5, 6, 7$). The relative distance between the other vehicles and the target vehicle is defined as Δd_2, Δd_3, Δd_4, Δd_5, Δd_6, and Δd_7, and the relative velocity is represented by $\Delta v_2 \triangleq v_{n2} - v_{n1}$, $\Delta v_3 \triangleq v_{n1} - v_{n3}$, $\Delta v_4 \triangleq v_{n4} - v_{n1}$, $\Delta v_5 \triangleq v_{n1} - v_{n5}$, $\Delta v_6 \triangleq v_{n6} - v_{n1}$, and $\Delta v_7 \triangleq v_{n1} - v_{n7}$. The types of target vehicle and other vehicles are represented by F_1, F_2, F_3, F_4, F_5, F_6, and F_7, vehicle-type values $F_i \in \{\text{small}, \text{middle}, \text{large}\}$. The driver's propensity is represented by P_1, P_2, P_3, P_4, P_5, P_6, and P_7, and the driver's propensity can take on values $P_i \in \{\text{radical}, \text{common}, \text{conservative}\}$, where $i = 1, 2, 3, 4, 5, 6, 7$.

The lane-changing process can be divided into three stages, namely the decision-making, implementation, and adjustment phase. The driver, considering the interest and the impact on the security of the surrounding vehicle cluster situation on him/her, determines the optimal lane and decides whether to change lanes. A driver's lane-changing decision-making process is a complex thought process, and it is inseparable from perception, logic, judgment, and causal reasoning and cannot be isolated from decomposition and synthesis. On the basis of understanding a drivers' physiological and psychological characteristics and the traffic situation (vehicle, the vehicle sequential activity, and a driver's propensity), the driver's lane-change decision-making process has to be analyzed and the lane-change
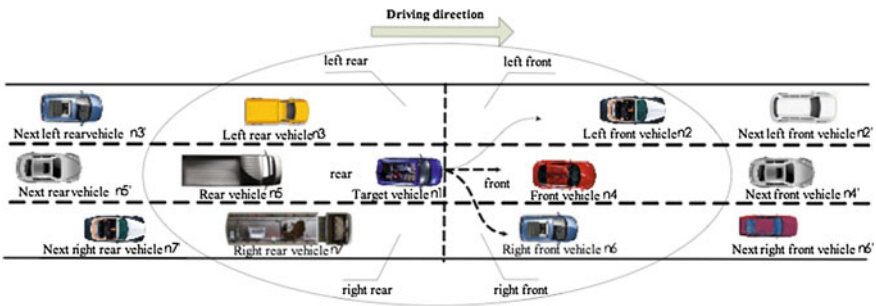
Fig. 1 Vehicle cluster situation when the target vehicle is located in the *middle line*

decision behavior characteristics should be analyzed with respect to human factors and the multiple perspectives of a given traffic situation. The consideration of these factors is presented in this paper.

Decision-Making Model of Lane Changing Based on Integrated Cognitive of Vehicle Cluster Situation

In the real world, the driver needs to consider various objective factors in a lane-changing decision-making process and judge whether to accept the front and rear acceptance gap. The fuzzy logic approach uses linguistic variables to approximate reasoning [7, 8], and it has been adapted to develop a subjective judgment process based on a driver's knowledge and experience. In this paper, a decision-making land-changing model that comprehensively considers the target vehicle driver's propensity and other aforementioned factors has been developed with this method.

Vehicle cluster situation exists with regard to the entire traffic environment around the target vehicle and changes with the dynamics of the transport element with respect to its surroundings. To facilitate further research into these dynamics, the concept of "force" in physics was borrowed to describe the effects of each subregion vehicle in a given interest-sensitive area on the target vehicle's lane-changing behavior. Here, if the vehicle in the subregion has a positive impact on the target vehicle to choose the lane where the subregion is located, then the dynamic that results in the vehicle staying in the subregion is defined as attraction; conversely, it is defined as repulsion.

The force of the subregion on the target vehicle is based on vehicle type, relative distance and relative velocity between the target vehicle, and the vehicle in the left front, left rear, front, rear, right front, and right rear subregions. This is followed with the calculation of the force of each lane on the target vehicle. Lastly, the target vehicle makes lane-changing decisions based on the force of each lane and target vehicle driver's propensity.

Fuzzy Variables and Corresponding Fuzzy Sets

(1) The relative distance between the target vehicle and surrounding vehicles is represented by Δd_i $(i = 2, 3, 4, 5, 6, 7)$

In the domain $\Delta d_i = \left\{ -d_i^0 \leq \Delta d_i \leq \Delta d_i^0 \right\}$ (wherein $-d_i^0$ is the lower bound of the domain, and Δd_i^0 is the upper bound of the domain), Δd_i possible fuzzy sets: $\{\mathrm{danger}, \mathrm{near}, \mathrm{medium}, \mathrm{far}\}$, the membership function of relative distance is shown in Fig. 2.
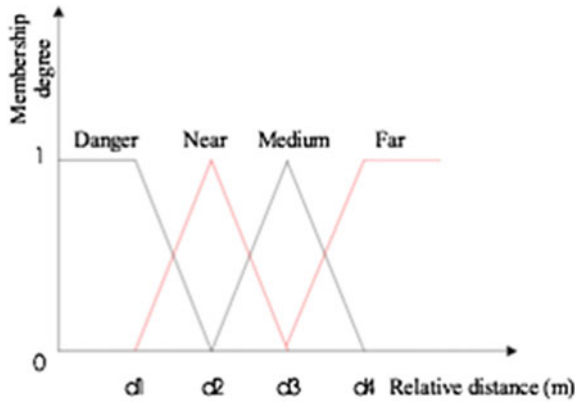
Fig. 2 Triangular membership function of relative distance
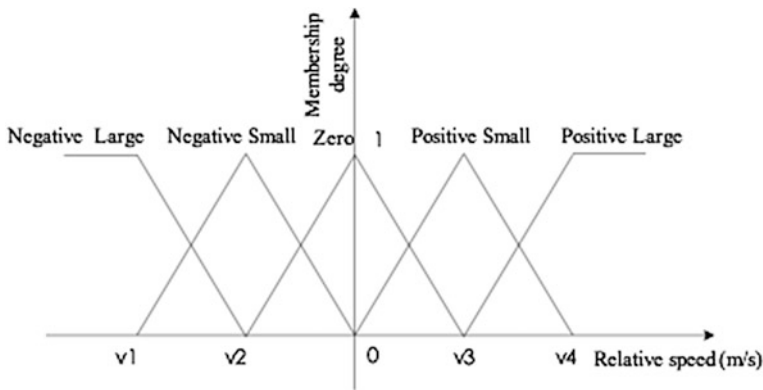
Fig. 3 Triangular membership function of relative speed

(2) The relative speed between the target vehicle and surrounding vehicles is represented by Δv_i ($i = 2, 3, 4, 5, 6, 7$)

In the domain $\Delta v_i = \left\{ -v_i^0 \leq \Delta v_i \leq \Delta v_i^0 \right\}$ (wherein $-v_i^0$ is the lower bound of the domain, and Δv_i^0 is the upper bound of the domain), Δv_i possible fuzzy sets: {Positive large, Positive small, Zero, Negative small, Negative large}, the membership function of relative speed is shown in Fig. 3.

d_1, d_2, d_3, and d_4 and v_1, v_2, v_3, and v_4 in the above two figures are the fuzzy subset boundary values of the relative distance domain and relative speed domain, respectively. A driver's relative distance experience changes with the change of relative speed. Therefore, the membership function of relative distance should contain relative speed parameters. The target vehicle and left rear vehicle of the

study subjects are used to discuss the calculation of fuzzy subset boundary values in the relative distance and the relative speed domains.

In Fig. 2, the given relative distance threshold value d_1 from "danger" to "near" is obtained by the following equation:

$$d_1 = \frac{v_{n3}^2 - v_{n1}^2}{2B_{max}} + \tau v_{n3} + \lambda_0 \tag{1}$$

B_{max} is the maximum acceleration of the vehicle, τ is the reaction time of the left rear vehicle, λ_0 is a parameter that needs to be determined. This threshold value is the minimum space between vehicles to avoid collision when the adjacent lane vehicle brakes and considering acceleration due to the occurrence of special events in front of the target vehicle and emergency braking in the target vehicle lane-change process. The relative distance threshold value from "medium" to "far" is

$$d_4 = D_0 + \frac{(v_{n3} - v_{n1})^2}{2a} + \lambda_1 \tag{2}$$

In which

$$D_0 = \frac{v_{n3}^2}{2B_{max}} + \tau v_{n3} + \lambda_2 \tag{3}$$

a is the acceleration of target vehicle, and λ_1, λ_2 are parameters that need to be determined. d_2 and d_3 are the medians of d_1 and d_4 in Fig. 2.

To ensure traffic safety, the relative speed should be within the appropriate range when the relative distance is in the boundary value "danger" d_1. Time to collision (TTC) can be used as an evaluation index of the risk of collision when the target vehicle speed is less than the left rear vehicle speed. It can be obtained by applying the following formula:

$$\text{TTC} = \left| \frac{\Delta d_3}{\Delta v_3} \right| \tag{4}$$

According to driver's lane-changing data and subjectively acceptable safety limits from literature [9, 10], the 5% quantile (2.6 s) and 25% quantile (5 s) of TTC statistics are used as classification boundary values of lane-change security. Thus, the relative distance threshold value v_1 from "Positive large" to "Positive small" in Fig. 3 is obtained by the following equation: $v_1 = -\frac{d_1}{2.6}$, the relative distance threshold value v_2 from "Positive small" to "Zero" is obtained by the following equation: $v_2 = -\frac{d_1}{5}$. v_3 and v_4 are opposites of v_2 and v_1 in Fig. 3.

Lane-Changing Decision-Making Based on Integrated Cognitive Vehicle Cluster Situation

The force of the vehicle in the left front, left rear, front, rear, right front, and right rear subregions to the target vehicle can be abstracted as attractions and repulsions (i.e., integrity for attraction and negative for repulsion). The magnitude of the force uses the effect size to describe: the greatest effect size of repulsion is represented by −1, while the greatest effect size of attraction is represented by 1 (Table 1).

Comprehensive analysis of a large volume of experimental data and simulations involving debugging were repeatedly conducted. The vehicle types as well as the consideration of relative distance and relative velocity between the target vehicle and the subregion vehicles in a given interest-sensitive area were included. A fuzzy logic approach was used to evaluate whether the determined effect size was reasonable.

Table 1 Effect size corresponding to different forces

Force	Strong repulsion	Medium repulsion	Weak repulsion	Zero	Weak attraction	Medium attraction	Strong attraction
Effect size	[−1, −0.8]	[−0.7, −0.4]	[−0.3, −0.1]	0	[0.1, 0.3]	[0.4, 0.7]	[0.8, 1]

Table 2 Fuzzy inference rules of effect size

Rule numbers	Target vehicle type	Left rear vehicle type	Relative distance	Relative speed	Effect size
1–4	Small	Small	Far (or medium)	Positive large (or positive small)	1
5–8	Midsize	Small	Far (or medium)	Positive large (or positive small)	1
9–12	Small	Midsize	Far (or medium)	Positive large (or positive small)	0.9
13–16	Midsize	Midsize	Far (or medium)	Positive large (or positive small)	0.9
17–24	Small (or midsize)	Small (or midsize)	Far (or medium)	Zero	0.8
25–28	Large	Small	Far (or medium)	Positive large (or positive small)	0.7
29–32	Small	Large	Far (or medium)	Positive large (or positive small)	0.7
33–36	Large	Midsize	Far (or medium)	Positive large (or positive small)	0.6

Table 2 (continued)

Rule numbers	Target vehicle type	Left rear vehicle type	Relative distance	Relative speed	Effect size
37–40	Midsize	Large	Far (or medium)	Positive large (or positive small)	0.6
41–44	Large	Large	Far (or medium)	Positive large (or positive small)	0.5
45–52	Small (or midsize)	Small (or midsize)	Near	Positive large (or positive small)	0.5
53–56	Large	Small (or midsize)	Near	Positive large (or positive small)	0.4
57–60	Small (or midsize)	Large	Near	Positive large (or positive small)	0.4
61–62	Large	Large	Near	Positive large (or positive small)	0.3
63–66	Large	Small (or midsize)	Far (or medium)	Zero	0.3
67–70	Small (or midsize)	Large	Far (or medium)	Zero	0.3
71–72	Large	Large	Far (or medium)	Zero	0.2
73–76	Small (or midsize)	Small (or midsize)	Near	Zero	0.2
77–80	Small (or midsize)	Small (or midsize)	Far	Negative small	0.2
81–82	Large	Small (or midsize)	Near	Zero	0.1
83–84	Small (or midsize)	Large	Near	Zero	0.1
85–86	Large	Small (or midsize)	Far	Negative small	0.1
87–88	Small (or midsize)	Large	Far	Negative small	0.1
89	Large	Large	Near	Zero	0
90	Large	Large	Far	Negative small	0
91–94	Small (or midsize)	Small	Medium (or near)	Negative small	0
95–98	Small (or midsize)	Midsize	Medium (or near)	Negative small	0
99–102	Large	Small (or midsize)	Medium (or near)	Negative small	−0.1
103–106	Small (or midsize)	Large	Medium (or near)	Negative small	−0.1
107–108	Large	Large	Medium (or near)	Negative small	−0.2
109–112	Small	Small (or midsize)	Danger	Positive large (or positive small)	−0.2
113–116	Midsize	Small (or midsize)	Danger	Positive large (or positive small)	−0.3

Table 2 (continued)

Rule numbers	Target vehicle type	Left rear vehicle type	Relative distance	Relative speed	Effect size
117–118	Large	Small	Danger	Positive large (or positive small)	−0.3
119–120	Small	Large	Danger	Positive large (or positive small)	−0.3
121–122	Large	Midsize	Danger	Positive large (or positive small)	−0.4
123–126	Midsize (or large)	Large	Danger	Positive large (or positive small)	−0.4
127–130	Small (or midsize)	Small	Danger	Zero (or negative small)	−0.4
131–134	Small (or midsize)	Midsize	Danger	Zero (or negative small)	−0.5
135–136	Large	Small	Danger	Zero (or negative small)	−0.5
137–138	Small	Large	Danger	Zero (or negative small)	−0.5
139–140	Midsize	Large	Danger	Zero (or negative small)	−0.6
141–144	Large	Midsize (or large)	Danger	Zero (or negative small)	−0.6
145–148	Small (or midsize)	Small	Far (or medium)	Negative large	−0.6
149–152	Small (or midsize)	Midsize	Far (or medium)	Negative large	−0.7
153–156	Large	Small (or midsize)	Far (or medium)	Negative large	−0.7
157–160	Small (or midsize)	Large	Far (or medium)	Negative large	−0.8
161–162	Large	Large	Far (or medium)	Negative large	−0.8
163–164	Small	Small	Near (or danger)	Negative large	−0.8
165–166	Small	Midsize	Near (or danger)	Negative large	−0.9
167–170	Midsize	Small (or midsize)	Near (or danger)	Negative large	−0.9
171–172	Large	Small	Near (or danger)	Negative large	−0.9
173–174	Small	Large	Near (or danger)	Negative large	−1
175–176	Large	Midsize	Near (or danger)	Negative large	−1
177–180	Midsize (or large)	Large	Near (or danger)	Negative large	−1

Taking the vehicle's effect size computation of the left rear side area on the target vehicle as an example, the fuzzy inference rules of effect size are shown in Table 2. One typical fuzzy rule describes the following:

If the target vehicle is a small car, the left rear vehicle is a small car, the relative distance between the target vehicle and the left rear vehicle is far, and the relative speed between the target vehicle and the left rear vehicle is positive large, then effect size of the left rear vehicle on the target vehicle is 1.

There are 180 rules included in Table 2, wherein the vehicle type, relative distance, and relative speed are input variables and the effect size is an output variable. The strength of the force from the subregion on the target vehicle can be obtained according to the effect size force. Calculating the force from the front and rear area vehicles area, the force of the lane on the target vehicle can be obtained by the fuzzy inference rules described in Table 3. Similarly, the force of the left lane and the right lane on the target vehicle can be obtained.

There are 49 rules that are presented in Table 3. The forces from front and rear area vehicles on target vehicle are input variables, and the force from the lane is an output variable. By considering the force from three lanes and the driver's propensity, the target vehicle driver's lane-change decision outcome can be inferred with the application of the fuzzy inference rules shown in Table 4.

There are 81 rules shown in Table 4. The forces from left, middle, and right lane on target vehicle and the target vehicle driver's propensity are input variables, and the decision to change lanes is an output variable. The instantaneous state of the vehicle on the road can be used to determine whether the target vehicle changes the lanes.

Table 3 Fuzzy inference rules of the force of the lane

Rule numbers	The force of the front area vehicles to the target vehicle	The force of the rear area vehicles to the target vehicle	The force of the lane to the target vehicle
1–4	Strong repulsion or medium repulsion	Strong repulsion or medium repulsion	Repulsion
5–10	Strong repulsion or medium repulsion	Weak repulsion or zero or weak attraction	Repulsion
11–14	Strong repulsion or medium repulsion	Medium attraction or st rong attraction	Zero
15–20	Weak repulsion or zero or weak attraction	Strong repulsion or medium repulsion	Repulsion
21–29	Weak repulsion or zero or weak attraction	Weak repulsion or zero or weak attraction	Zero
30–35	Weak repulsion or zero or weak attraction	Medium attraction or strong attraction	Attraction
36–39	Medium attraction or strong attraction	Strong repulsion or medium repulsion	Zero
40–45	Medium attraction or strong attraction	Weak repulsion or zero or weak attraction	Attraction
46–49	Medium attraction or strong attraction	Medium attraction or strong attraction	Attraction

Table 4 Lane change decision-making fuzzy inference rules

Rule numbers	Left lane force	Middle lane force	Right lane Force	Driver's propensity	Lane change decision
1–3	Repulsion	Repulsion	Repulsion	All	Lane keeping
4–5	Repulsion	Repulsion	Zero	Radical type or common type	Right lane changing
6	Repulsion	Repulsion	Zero	Conservative type	Lane keeping
7–9	Repulsion	Repulsion	Attraction	All	Right lane changing
10–12	Repulsion	Zero	Repulsion	All	Lane keeping
13	Repulsion	Zero	Zero	Radical type	Right lane changing
14–15	Repulsion	Zero	Zero	Common or conservative type	Lane keeping
16–18	Repulsion	Zero	Attraction	All	Right lane changing
19–21	Repulsion	Attraction	Repulsion	All	Lane keeping
22–24	Repulsion	Attraction	Zero	All	Lane keeping
25	Repulsion	Attraction	Attraction	Radical type	Right lane changing
26–27	Repulsion	Attraction	Attraction	Common or conservative type	Lane keeping
29–30	Zero	Repulsion	Repulsion	All	Left lane changing
31	Zero	Repulsion	Zero	Radical type	Left lane changing
32	Zero	Repulsion	Zero	Common type	Lane keeping
33	Zero	Repulsion	Zero	Conservative type	Lane keeping
34–36	Zero	Repulsion	Attraction	All	Right lane changing
37–39	Zero	Zero	Repulsion	All	Lane keeping
40–41	Zero	Zero	Zero	Common or conservative type	Lane keeping
42	Zero	Zero	Zero	Radical type	Left lane changing
43–45	Zero	Zero	Attraction	All	Right lane changing
46–48	Zero	Attraction	Repulsion	All	Lane keeping
49–51	Zero	Attraction	Zero	All	Lane keeping
52	Zero	Attraction	Attraction	Radical type	Right lane changing
53–54	Zero	Attraction	Attraction	Common or conservative type	Lane keeping
55–57	Attraction	Repulsion	Repulsion	All	Left lane changing

(continued)

Table 4 (continued)

Rule numbers	Left lane force	Middle lane force	Right lane Force	Driver's propensity	Lane change decision
58–60	Attraction	Repulsion	Zero	All	Left lane changing
61	Attraction	Repulsion	Attraction	Radical type	Left lane changing
62–63	Attraction	Repulsion	Attraction	Common or conservative type	Right lane changing
64–66	Attraction	Zero	Repulsion	All	Left lane changing
67	Attraction	Zero	Zero	Radical type	Left lane changing
68–69	Attraction	Zero	Zero	Common or conservative type	Lane keeping
70	Attraction	Zero	Attraction	Radical type	Left lane changing
71–72	Attraction	Zero	Attraction	Common or conservative type	Lane keeping
73	Attraction	Attraction	Repulsion	Radical type	Left lane changing
74–75	Attraction	Attraction	Repulsion	Common or conservative type	Lane keeping
76	Attraction	Attraction	Zero	Radical type	Left lane changing
77–78	Attraction	Attraction	Zero	Common or conservative type	Lane keeping
79	Attraction	Attraction	Attraction	Radical type	Left lane changing
80–81	Attraction	Attraction	Attraction	Common or conservative type	Lane keeping

Data Collection and Model Validation

Identification of the Driver's Propensity

In reference to the questionnaire and the identification of the method to determine a driver's propensity, (i.e., in Ref. [11]), 50 drivers' propensities were identified.

Actual Vehicle Verification

In actual vehicle experiments, time series of measured data, which affect a driver's lane-changing behavior were gathered.

(1) **Experimental Equipment**

Given a city road environment, a dynamic driver–vehicle–environment information acquisition system (including PsyLAB human factors engineering experiment wireless sensors, Bioharhess Bioharness 3.0 portable physiological signal measurement instrument, Tobii Glasses eye tracker, high-definition cameras, WZC-2D steering parameter tester, WTC-1 pedal force control force meter, SG299-GPS non-contact multi-function speedometer, CTM-8A non-contact multi-function speedometer, BTM300-905-200 laser range finder sensor, X5000 vehicle data recorders, HDTV Camera, Notebook Computer, etc., as shown in Fig. 4) was used to collect and process the collected data. In addition, PsyLAB software, AcqKnowledge 4.2.1, Tobii Studio, Minivcap monitoring system software, SPSS17.0, Ulead video studio 10, etc., were used.

(2) **Experimental Conditions**

On a morning of a workday, between 7:30 am and 9:30 am, with a good weather conditions, Beijing Road–Xincun xi Road–Yuanshan Road–Huaguang Road in Zibo City (the length is approximately 9 km, as shown in Fig. 5) was selected. This location was defined as an experimental non-free flow situation, where 50 drivers
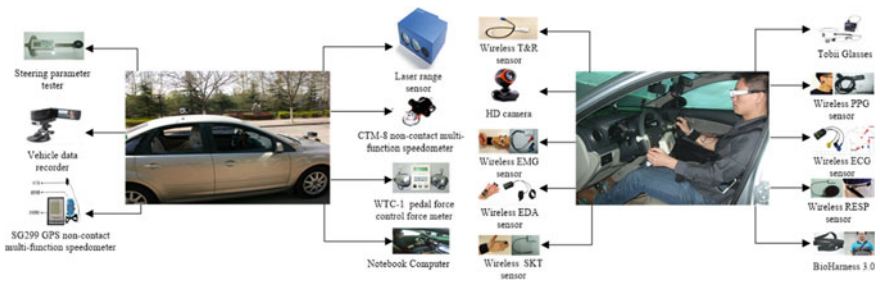
Fig. 4 Schematic of dynamic driver–vehicle–environment information acquisition systems

Fig. 5 Proposed experimental route

with different driving propensities operated small, medium, and large vehicles, which were loaded with dynamic environment information collection systems. These vehicles traveled along the test route, and vehicle traveling data were collected in the process.

(3) **Experiment Content**

In an actual vehicle experiment, the target vehicle has been calibrated with a GPS multi-functional instrument, laser ranging sensor, and video acquisition system. The calibration is performed by using a non-experimental design contact dynamic vehicle environment with a comprehensive information acquisition system to organize and design the experiment. Additionally, the target vehicle needs to run on the specified route to ensure the normal operation of the aforementioned instruments in the experimental process. Dynamic comprehensive information collection systems can then provide a real-time record of the relevant experimental data.

(4) **Data and Analysis**

The experimental data are shown in Table 5.

Model Calibration

Experimental data describing the lane-change behavior of different driver types were acquired. Based on the reference value, calibration of the parameters was completed with real vehicle test data. Calibration parameters are shown in Table 6.

Table 5 Experimental data that can be collected

Data	Relative distance between target vehicle and the surrounding vehicles (m)						Relative speed between target vehicle and the surrounding vehicles (m/s)					
	Left front	Left rear	Front	Rear	Right front	Right rear	Left front	Left rear	Front	Rear	Right front	Right rear
Code	Δd_2	Δd_3	Δd_4	Δd_5	Δd_6	Δd_7	Δv_2	Δv_3	Δv_4	Δv_5	Δv_6	Δv_7

Table 6 Model calibration of lane-changing decision-making model

Parameters		v_1	v_2	v_3	v_4	d_1	d_2	d_3	d_4
Value	Radical	−3.3	−1.8	1.8	3.3	8.5	22.5	36.3	51.1
	Common	−4.7	−2.5	2.5	4.7	12.3	25.4	40.1	60.2
	Conservative	−8.6	−4.6	4.6	8.6	22.6	35.4	50.4	71.1

Model Verification

(1) **Simulation Experiment**

Simulation experiments are carried out with a simulator made by the Japan FORUM8 enterprise. The simulation software used in the driving simulator is UC-win/Road (shown in Fig. 6). Simulated three-lane driving scenes should be developed via experiments. Drivers should be trained to use the simulator before experiments are conducted. To prevent the interference of drivers in the experimental process and to store experimental data afterward, the method described in Ref. [11] was applied, and drivers' propensity types were accordingly distinguished. Lane-change results were predicted based on the acquired data, and the results are shown in Table 7.

Based on Table 7, the lane-changing decision-making model was developed to present prediction data, which is in high agreement with results obtained in actual test situations.
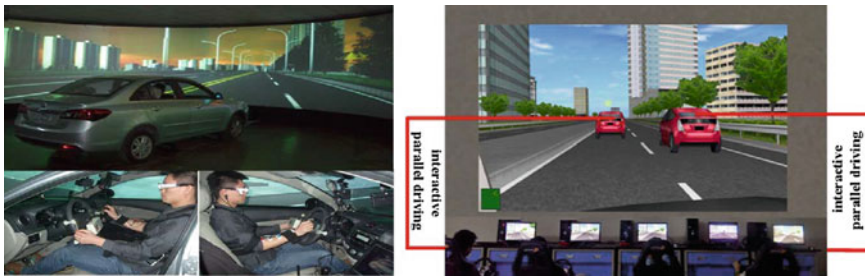
Fig. 6 Interactive parallel-driving simulated experiments

Table 7 Imitation results

Driver's number	Predict times	The comparison of prediction and recognition results		Accuracy (%)
		Agree times	Disagree times	
1	80	68	12	85.0
2	80	69	11	86.3
3	80	65	15	81.3
4	80	69	11	86.3
5	80	66	14	82.5
6	80	70	10	87.5
7	80	64	16	80.0
8	80	67	13	83.8
Mean value	80	67.25	12.75	84.06
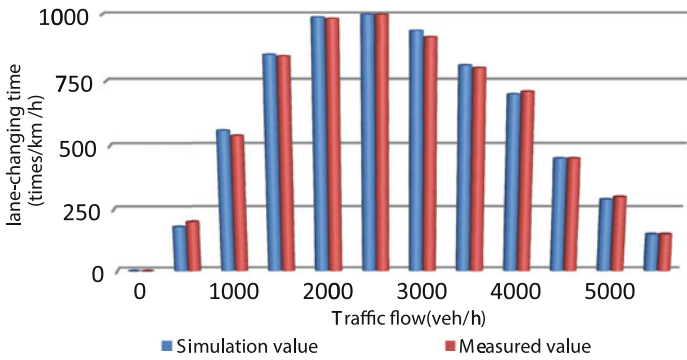
Fig. 7 Lane-changing time comparison of simulation and actual measured values
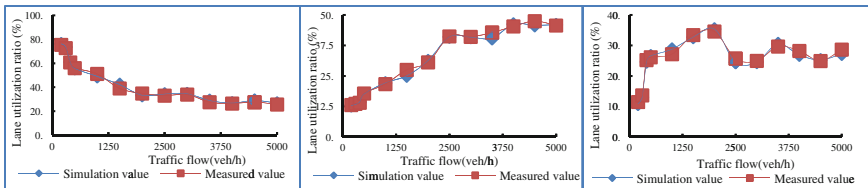
Fig. 8 Lane utilization ratio comparison of simulation and actual measured value in *right*, *middle*, and *left* lane

(2) **Simulation Verification**

The model established in this paper was used for simulations; a lane-change time and lane-occupancy-rate comparison chart between actual measured value and simulation value is shown in Figs. 7 and 8.

Conclusion

Based on a comprehensive analysis of lane-changing behavior and the consideration of the impact of the target vehicle and surrounding vehicle cluster situation, it was possible to task concentration comprehensive cognitive reaction on the process of lane change. To accomplish this, the development of a new, comprehensive lane-changing decision-making model that considers a target vehicle driver's propensity and the surrounding vehicles' types, distance, and speed in complex vehicle cluster environments was described in this paper. Field data, virtual reality data, and simulation data used for validation support the claim that this model can provide theory-based insight into driver's behavior for the in-depth study of microscopic traffic flow simulation. It should be noted that the lane-changing

decision behavior will be different under different traffic conditions. In addition, different drivers' decision-making behavior may also be different under the same traffic lane-changing environments. Therefore, the model parameters need to be calibrated with respect to the particular drivers and the corresponding traffic environment for more accurately reflecting the driver's lane-changing decision behavior.

References

1. McDonald, Mike, Jianping Wu, and Mark Brackstone. 1997. Development of a fuzzy logic based microscopic motorway simulation model. In *Proceedings of the IEEE Conference on Intelligent Transportation Systems, Boston, 1997*, 82–87.
2. Wu, Jianping, Mark Brackstone, and Mike McDonald. 2000. Fuzzy sets and systems for a motorway microscopic simulation model. *Fuzzy Sets and Systems* 116: 65–76.
3. Das, S., and B.A. Bowles. Simulations of highway chaos using fuzzy logic. In *Fuzzy Information Processing Society (NAFIPS), New York, 1999*, 130–133.
4. Errampalli, Madhu, Masashi Okushima, and Takamasa Akiyama. 2008. Fuzzy logic based lane change model for microscopic traffic flow simulation. *Journal of Advanced Computational Intelligence and Intelligent Informatics* 12 (2): 172–181.
5. Moridpour, Sara, Geoff Rose, and Majid Sarvi. 2009. Modelling the heavy vehicle drivers' lane changing decision under heavy traffic conditions. *Road & Transport Research: A Journal of Australian and New Zealand Research and Practice* 18 (4): 49–57.
6. Wu, Lei, Xiaohui Wang, Xinyue Yang, and Xiaoyuan Wang. 2007. Study on the recognition of traffic situation and its state transition mechanism. *Communications Standardization* (2/3): 61–65.
7. Karatop, Buket, Cemalettin Kubat, and Özer Uygun. 2015. Talent management in manufacturing system using fuzzy logic approach. *Computers & Industrial Engineering* 86: 127–136.
8. Petrović, Dejan V., Miloš Tanasijević, Vitomir Milić, Nikola Lilić, Saša Stojadinović, and Igor Svrkota. 2014. Risk assessment model of mining equipment failure based on fuzzy logic. *Expert Systems with Applications* 41 (18): 8157–8164.
9. Minderhoud, Michiel M., and Piet H.L. Bovy. 2001. Extended time-to-collision measures for road traffic safety assessment. *Accident Analysis and Prevention* 33(1): 89–97.
10. Lee, S.E., E.C.B. Olsen, and W.W. Wierwille. 2004. *A Comprehensive Examination of Naturalistic Lane-Changes*. Washington, DC: National Highway Traffic Safety Administration.
11. Wang, Xiaoyuan, Jin Liu, and Jinglei Zhang. 2012. Dynamic recognition model of driver's propensity under multilane traffic environments. *Discrete Dynamics in Nature and Society* 2012: 1–15.

Novel Design of Head-Up Display System Based on Safety Control

Rongjie Lin, Qian Cheng, Xiaobei Jiang and Wuhong Wang

Abstract As a part of driver assistance system, head-up display system can improve the performance and safety of driving activity. Questionnaires are conducted to understand the drivers' degree of attention on various driving information and to extract key parameters that affect the driving behaviour. The workflow and architecture of the head-up display system based on safety control are estimated. In the system, smart terminal plays an important role, which can help lower the expense of the hardware and enrich the information sources. Interface design of the HUD system for the road segment and intersection is elaborated.

Keywords Head-up display · Safety control · Interface design · Architecture

Introduction

Researchers have always tried to improve the driving safety. As the number of vehicles grows, traffic accidents have long been a depressing social–economical problem that has caused increasing concern to the public worldwide [1]. Statistical data shows that 80% of the traffic accidents are caused by the driver's factor. According to the causation analysis of motor accidents from countries over the world, the precipitating factor of 55–60% accidents caused by drivers are the visual errors, and 35–40% accidents are caused by the decision-making errors [2]. Consequently, it is vital to enhance the visual and decision-making stage to improve the drivers' safety.

R. Lin · Q. Cheng · W. Wang (✉)
Beijing Institute of Technology, Beijing, China
e-mail: wangwuhong@bit.edu.cn

X. Jiang (✉)
Beijing Institute of Technology, Beijing, China
e-mail: jiangxiaobei@bit.edu.cn

X. Jiang
Technical University of Munich, Munich, Germany

Due to the limitation of the individual driver's capability, a great quantity of attention has been paid to the driver assistance technology, which can help strengthen the drivers' capability. Driver assistance technology consisting of passive and active safety fields tries to simplify the operating of the driving behaviour and make it possible for drivers to draw more attention to the traffic. Mechanical assistances such as gasoline, automatic transmission, brake and steering booster are applied in vehicles in the early time. In the beginning of the 1970s, the antilock braking system (ABS) was presented firstly by Mercedes-Benz, which marked the beginning of electronic assistance. The electronic stability program applied by Daimler-Chrysler, automatic cruise control system, lane keeping assistant system, head-up display (HUD in short below) system and some other assistance systems were gradually invented and equipped in vehicles [3]. Previous relevant studies reveal that drivers mostly use (over 90%) their vision in obtaining related driving information [4]. Considering that visual errors are important causation among precipitating factors of driver accidents, scientists are trying to explore methods to assist drivers to enhance the visual capability. The HUD system, which is one aspect of driver assistance technology, has gradually become a research focus. The first practical application of HUD was in the Hawker-Siddeley Buccanneer in 1960 [5]. HUD system has been put in use on aircrafts universally at the end of 1970s. The automobile manufacturer General Motor firstly fixed the HUD system on vehicles in 1988. Manufacturers have then gradually produced vehicles with HUD.

HUD system has gained much attention of scientists and engineers since its first appearance. Previous studies presented many benefits of HUD use. It was found by researchers that the automotive HUD system provides the advantages of keeping drivers' eyes near the road and minimizes the focal re-accommodation time by displaying the information needed on the windshield in front of the driver [5–7]. R. J. Kiefer and N.A. Kaptein proposed that HUD can reduce the frequency and duration of the driver's eyes-off-the-road. With the assistance of HUD, drivers can steer easily and respond to the information quickly to guarantee safety and stability [8, 9]. Y.C Liu and V. Charissis et al. found out that drivers' response time to an urgent event is faster with a HUD than a HDD (head-down display), and speed control is more consistent with HUD [10–12]. Y.C Liu's [13] research also revealed that using HUD causes less mental stress for drivers and is easier for first time users to use. Moreover, M. Tonnis et al. found that most drivers feel safer when driving with an HUD [14]. In consequence, HUD system is expected to become a vital device for most vehicles in the future. Given the advantages of HUD, it seems to be a feasible and superior alternative or auxiliary visual display interface on vehicles [15]. However, two potential negative impacts of HUD use on driving safety were found by D.R. Tufano. Tufano indicated that HUD focal distance may affect drivers' accommodation and perception of actual objects when driving, while HUD images may clutter or block drivers' view and affect visual attention [16]. Y.C Liu found that the driver's enhanced awareness and the cognitive capture might affect driving performance and aggravate the driver's workload. Meanwhile, the so-called novelty effect for using new technology product was found [4].
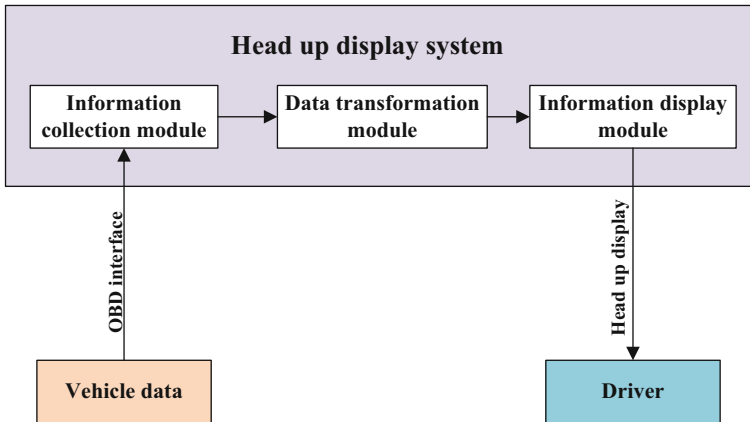
Fig. 1 Workflow of traditional HUD system

Traditional HUD systems can be divided into three working modules, which, respectively, are information collection module, data transformation module and information display module. Its workflow is shown in Fig. 1. The main function of traditional HUD systems is to realize the head-up display of the vehicle data. With the development of the smart terminal and 4G technology, traditional HUD system cannot meet the safety needs of drivers. New system architecture and workflow should be established to promote the popularization and development of HUD.

This paper proposes the novel design of HUD system based on safety control. Demand investigation is conducted to analyse the demand of drivers about the system. According to the investigation and analysis of the intention to use head-up display system for the driver, driver's intention and behavioural characteristics are summarized. Then the new architecture and workflow of HUD system are established. Each layer of the system architecture is elaborated. Meanwhile, interface design of straight road and intersection based on the intelligent terminal and of intersection based on the intelligent windshield are elaborated, so that it can be the theoretical basis for further applications.

Methodology and Results

To understand the drivers' degree of attention on the various driving information and to extract key parameters that affect driving behaviour, this paper adopted questionnaires to analyse the characteristics of driving behaviour. We issued the survey questionnaire at Weigongcun, Zhaogongkou and Beijing South Railway Station to ensure that the survey data was representative, and the respondents sought a uniform distribution of driving experience. The total number of valid collected questionnaire is 147. Among the respondents, the percentage of drivers

whose driving experience is less than one year, one to three years, three to six years, more than six years, respectively, is 19.7, 27.2, 25.9 and 27.2%. In the questionnaire, statistical processing, frequency analysis, descriptive statistical analysis, variance analysis and cross-table analysis function of SPSS software are applied to analyse thoroughly the understanding, acceptance and preference of drivers on the head-up display system equipped on the vehicle.

The statistical data of the questionnaires shows 69.4% of the drivers are used to lower their head to check the dashboard while 66.7% of the drivers consider checking the dashboard is dangerous, which may cause an accident. 63.3% of the drivers have never heard of the HUD system, and just 4.1% of the drivers have ever used the system. Meanwhile, 87.8% of the drivers consider it helpful to safe driving when more driving information is collected. 70.1% of the drivers hope that the driving information can be projected on the windshield. The investigation result shows that although popularizing rate of the HUD system is low, most drivers have positive attitude on the system which implies high demand for HUD system.

Respondents evaluated the degree of concern of the main driving information, such as remaining oil, tyre pressure, engine speed, navigation message, turn signal, engine failure, speed of front vehicle, traffic report, vehicle speed and space headway. The Richter magnitude scale is established to describe the degree of concern, in which 1–5, respectively, represent refuse to pay attention, less concerned, no matter, more concerned and very concerned. According to the principle of statistics, arithmetic mean is applied for further analysis. The arithmetic mean of the degree of concern of driving information is ranked, as in the Fig. 2, which
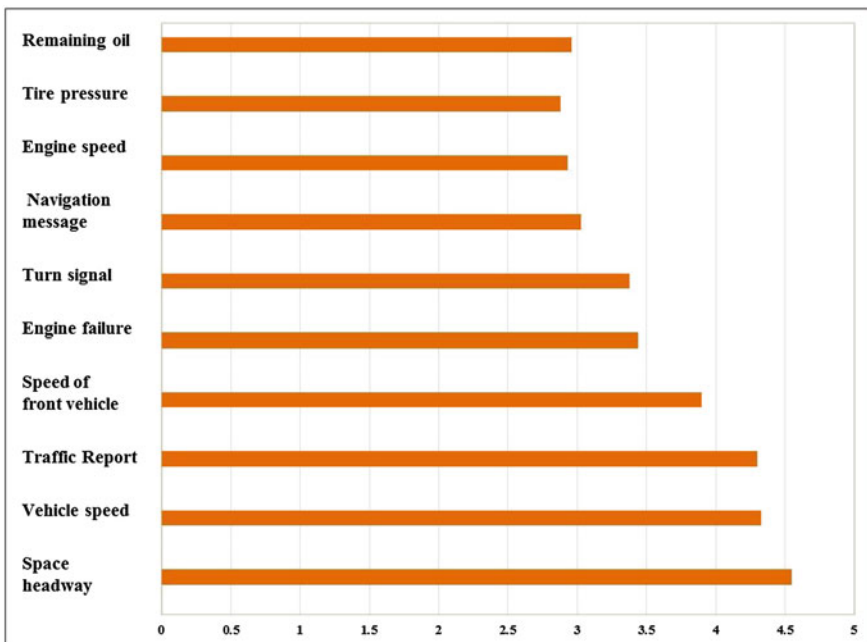
Fig. 2 The degree of concern for the driver of the main driving information

shows that speed of front vehicle, traffic report, vehicle speed and space headway are the information that drivers are more concerned, especially the space headway. The results can be applied to design the function of the HUD systems and the interface design.

Architecture Design of the HUD System

The workflow of the HUD system based on safety control is different from the traditional system, which is shown in Fig. 3. As the core of the workflow, safety control module plays an important role in improving the driving safety. The smart terminal such as the mobile phone or tablet PC is applied as a basis of the system with the objective to lower the expense of the hardware and enrich the information sources. With the smart terminal, the display mode can be more flexible. In the new workflow, the smart terminal can match the OBD transmission interface by Bluetooth, so that the vehicle data and the sensor data can be transferred to information collection module of the smart terminal wirelessly. Besides, GPS data and wireless data can be transferred to information collection module as well by the built-in module of the terminal. So the information collection module can collect abundant information from various sources and transfer the information to the process module for data conversion. Then, safety control module will analyse the information with gap acceptance theory, security zone theory and safety control algorithms so that the safety critical point and dangerous driving status can be identified. After the identification, the image or sound warning signal will be sent to the information display module and eventually showed to drivers. Equipped with the novel HUD system based on safety control, it can be easy for drivers to identify the dangerous driving status and correct their driving error to keep safe.

According to the HUD function and connections of HUD elements, the HUD system based on safety control is divided into four layers, namely function layer, information display layer, safety control layer, and information collection and process layer. The architecture of the system is shown in Fig. 4. The function layer
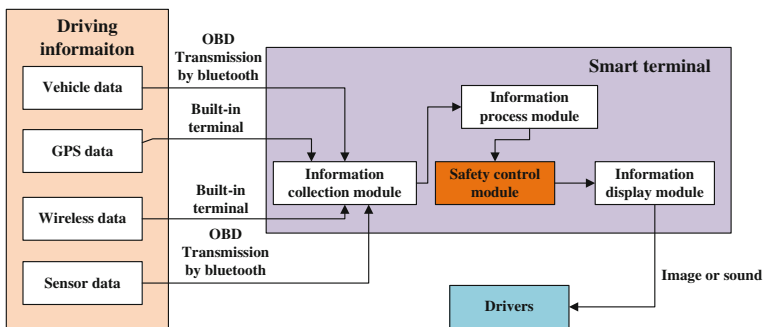
Fig. 3 The workflow of the HUD system based on safety control
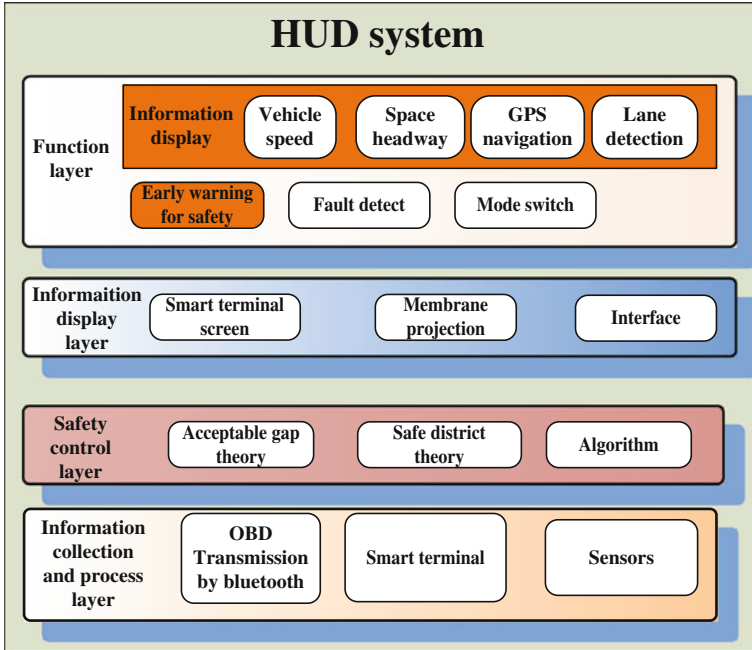
Fig. 4 The architecture of the HUD system based on safety control

contains system functions such as information displaying, early warning for safety, fault detecting and mode switching. The information on vehicle speed, space headway, GPS navigation and lane detection can be projected on the windshield. The information display layer consists of smart terminal screen, membrane projection and human–machine interface. The key layer is the safety control layer which contains acceptable gap theory, safe district theory and safety control algorithm. The traditional information collection layer and process layer are merged in the architecture because of the application of smart terminal. This layer includes OBD transmission by Bluetooth, smart terminal and sensors. The four layers are important components of the HUD system, which are interrelated and mutually influenced. Technical breakthrough of each layer will promote the development of HUD system effectively.

Compared with the traditional HUD system, the novel design of the system based on safety control has three advantages. Firstly, safety control layer is replenished to the architecture so that the function of safety control and early warning are strengthened. The safety control layer can help drivers quickly identify the dangerous driving status to keep safe and reliable. Secondly, based on the smart terminal, the novel architecture can take full advantage of the smart terminal so that the hardware expense of the system can be reduced, and the installation of the system can be easier. Thirdly, the usage of Bluetooth can realize the wireless transmission.

Interface Design of the HUD System

In dangerous situations, such as low visibility navigation in a highway environment, interface delays or cognitive capture effect due to an ineffective design can be fatal. In general, interface design should follow a human-centred approach [12]. Well-designed interface can help drivers collect the information they need and keep driving safe. Considering the cognitive capture effect and difference between the road segment and intersection, interface design is divided into two situations. On the road segment, the traffic environment is relatively simple, so that the information offered to the driver can be more succinct than the complicated intersection environment. This paper proposes the interface design based on the smart terminal. Tablet PC of certain band is chosen for the design whose screen size is 262 × 180 mm.

At the road segment, the information showed by the interface of HUD system is mainly vehicle speed, lane assist and space headway. The interface design in the situation is shown in Fig. 6. Here, 60 km/h represents the vehicle speed while the red box represents the space headway. The red box symbol is more readable than number. The box is coloured as red, yellow and green. When the space headway is safe, it is red. It will turn to yellow when the space headway approaches the safety critical point. If the space headway exceeds the critical point, the box is red and will give the driver early warning. Succinct but important information is delivered to drivers by the interface, and drivers can understand easily and quickly (Fig. 5).

The interface design at the intersection is shown in Fig. 6. The information is displayed in the interface such as vehicle speed, lane assist, space headway, turning

Fig. 5 Interface design at the road segment

Fig. 6 Interface design at the intersection

light status, navigation information and intersection information. The difference in segment situation is the added navigation information and intersection information. With the information displayed, drivers can quickly get the information of the intersection and traffic environment.

Conclusion

As a part of driver assistance system, HUD system can improve the performance and safety of driving. Questionnaires are conducted to understand the degree of drivers' attention on the various driving information and to extract key parameters that affect the driving behaviour. The workflow and architecture of the head-up display system based on safety control are estimated. In the system, smart terminal as a basis of the system plays an important role which can help lower the expense of the hardware and enrich the information sources. Interface design of the HUD system for the road segment and intersection is elaborated. The research can be the theoretical basis for further practical application. Besides, researches on factors such as driving simulation need to be conducted in the future.

References

1. Clare, K. 2003. WHO acts on road safety to reverse accident trends. *The Lancet* 362 (4): 1125.
2. Wang, W.H., Q. Cao, and D.M. Liu. 1994. Analysis and assessment of driver's operation reliability. *Vehicle Engineering* 16 (4): 207–213.
3. B. Heiner. 2011. Traffic safety through driver assistance and intelligence. *International Journal of Computational Intelligence Systems* 4 (3): 287–296.
4. Lansdown, T.C. 1997. Visual allocation and the availability of driver information. In *Traffic and transport psychology: Theory and application*, ed. T. Gothengater, and E. Carbonell, 215–223. Amsterdam: Pergamon.
5. Kabayashi, S.O., M. Sakata., J. Fukano., S. Daidoji., C. Hashimoto., and T. Ishikawa. 1989. Development of practical heads-up display for production vehicle application, SAE Technical Paper No. 890559, New York.
6. Weintraub, D.J., R.F. Haines., and R.J. Randle. 1984. The utility of head-up displays: Eye-focus vs. decision times. In *Proceedings of the 28th Human Factors and Ergonomics Society Annual Meeting*, 529–533. Santa Monica: Human Factors Society.
7. Weintraub, D., R. Haines., and R. Randle. 1985. Head-up display (HUD) utility II: Runway to HUD transitions monitoring eye focus and decisiontimes. In *Proceedings of the 29th Human Factors and ErgonomicsSociety Annual Meeting*, Human Factors Society, Santa Monica, 615–619.
8. Kiefer, R.J. 1991. *Effect of a head-up versus head-down digital speedometer on visual sampling behaviour and speed control performance during daytime automobile driving*, SAE Technical Report Paper No. 910111, New York.
9. Kaptein, N.A. 1994. *Benefits of in-car head-up displays*, TNO Report No. TNO-TM 1994 B-20, TNO Human Factors Research Institute, The Netherlands.
10. Liu, Y.C., and M.H. Wen. 2004. Comparison of head-up display (HUD) vs. head-down display (HDD): Driving performance of commercial vehicle operators in Taiwan. *International Journal of Human-Computer Studies* 61: 679–697.
11. Charissis, V., S. Arafat., W. Chan., and C. Christomanos. 2006. Driving simulator for head up display evaluation: driver's response time on accident simulation cases. In *Driving Simulation Conference*, DSC'06 Asia/Pacific, Tsukuba, Tokyo, Japan.
12. Charissis, V., and S. Papanastasiou. 2010. Human–machine collaboration through vehicle head up display interface. *Cognition, Technology & Work* 12: 41–50.
13. Liu, Y.C. 2003. Effect of using head-up display in automobile context on attention demand and driving performance. *Displays* 24: 157–165.
14. Tonnis, M., C. Lange., and G. Klinker. 2007. Visual longitudinal and lateral driving assistance in the head-up display of cars. In *Proceedings of the Sixth IEEE and ACM International Symposium on Mixed and Augmented Reality*, Nara, Japan, 128–131.
15. Liu, Y.C. 2003. Effects of using head-up display in automobile context on attention demand and driving performance. *Displays* 24: 157–165.
16. Tufano, D.R. 1997. Automotive HUDs: The overlooked safety issues. *Human Factors* 39: 303–311.

Prediction Model on Energy Consumption of Highway Transportation in Inner Mongolia Based on ARMA

Zhengyu Wang, Yueying Huo and Zhenyu Liu

Abstract Energy consumption of highway transportation in Inner Mongolia gradually increases, which seriously restricts sustainable economy development. So, this paper examines energy consumption of highway transportation prediction model using ARMA. Firstly, ARMA model is introduced. Then, based on the Unit Energy Consumption (energy consumption per unit Converted Transport Turnover) from 1999 to 2012, prediction model for Unit Energy Consumption is proposed using ARMA model. Finally, the Unit Energy Consumption in 2020 is forecasted according to the proposed prediction model. The study results of this paper provide useful evidences for drawing up the transportation 13th Five-year Plan and for formulating energy-saving and emission-reduction policy in Inner Mongolia.

Keywords Energy consumption · Highway transportation · Inner mongolia · ARMA model

Introduction

Inner Mongolia is located in north of China. The distance from west to east is 2400 km; the distance from south to north is 1700 km; and its area is 1.183 million square kilometers. There are rich natural resources in Inner Mongolia, for example prairie, forest, farmland areas per person rank the top in China and rare earth reservation occupies the primary position in the world. The economic development in Inner Mongolia is very fast, where GDP per capita exceeds the average level of Chinese mainland. Also in recently years, highway construction in Inner Mongolia increasingly enhanced, its highway mileage is 167,515 km until the end of 2013, and average annual growth rate is 19.7% [1–5].

Z. Wang · Y. Huo (✉) · Z. Liu
Transportation Institute of Inner Mongolia University, Hohhot 010070, Inner Mongolia, People's Republic of China
e-mail: huoyueying2008@163.com

With the quick economic development and increased highway mileage, highway transport volume (including passenger traffic and freight traffic) in Inner Mongolia is experiencing rapid increase and the average growth rate per annum is 8% [1–5]. In this context, energy consumptions of highway transport in Inner Mongolia also get very high. High energy consumption will seriously restrict sustainable economy development. Thus, it is very necessary to forecast energy consumption of highway transport. This article aims to develop a prediction model on energy consumptions of highway transport in Inner Mongolia and forecast the energy consumptions in 2020. The study results will provide useful evidences for drawing up the transportation 13th Five-year Plan and for formulating energy-saving and emission-reduction policy.

In this article, energy consumption per unit Converted Transport Turnover (short for "Unit Energy Consumption") is modeled and forecasted. Converted Transport Turnover (converted ton-km) refers to the sum of freight turnover (ton-km) and the converted freight turnover from passenger turnover (passenger-km).

This article is organized as follows. The second section describes the modeling method—ARMA model. Following which the modeling process is depicted for Unit Energy Consumption prediction model. Then, Unit Energy Consumption in 2020 is forecasted with the established prediction model. Finally, some discussions about the prediction model are presented.

ARMA Model

ARMA model is the short form of Autoregressive and Moving Average model. In the statistical analysis of time series, ARMA model provides a parsimonious description of a stationary stochastic process in terms of two polynomials, one for the autoregression and the second for the moving average [6]. The general ARMA model was developed in the 1951 and popularized in the 1971 [7]. Given a time series of data y_t, ARMA model is a tool for understanding and, perhaps, predicting future values in this series.

The Form of ARMA Model

ARMA model consists of two parts, an autoregressive (AR) part and a moving average (MA) part [8–10]:

1. AR(p)

   The notation AR(p) refers to the autoregressive model of order p. The AR(p) model is written as follows:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t \tag{1}$$

where y_t is a stationary time series, p is the order of autoregressive model, ϕ_i $(i = 1, 2, \ldots, p)$ are undetermined parameters, and ε_t is white noise (error).

2. $MA(q)$

   The notation $MA(q)$ refers to the moving average model of order q. The $MA(q)$ model is written as follows:

$$y_t = \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{2}$$

where y_t is a stationary time series, q is the order of the moving average model, θ_j $(j = 1, 2, \ldots, q)$ are undetermined parameters, and ε_t is white noise (error).

3. $ARMA(p, q)$

   The notation $ARMA(p, q)$ refers to the model with p autoregressive terms and q moving average terms. The model form is as follows:

$$y_t = \phi_1 y_{t-1} + \phi_2 y_{t-2} + \cdots + \phi_p y_{t-p} + \varepsilon_t - \theta_1 \varepsilon_{t-1} - \theta_2 \varepsilon_{t-2} - \cdots - \theta_q \varepsilon_{t-q} \tag{3}$$

where the symbols here have the same meaning as expressions (1) and (2).

The Modeling Process Based on ARMA Model

When we use ARMA model to develop a prediction model and further identify a prediction value, the general procedure involves four steps below [8–10]:

1. Stationarity test for time series.

   Before estimate values in time series, unit root test is conducted to identify the order of integration of time series, the purpose for which is to confirm the time series is stationary.

2. Model order determination

   Orders p and q in ARMA model need to be predetermined. The values for p and q can be preliminarily determined though observing autocorrelation coefficients and partial autocorrelation coefficients. Then, they are eventually determined based on Akaike Information Criterion (AIC). The specific principles for determining p and q are shown in Table 1.

Table 1 Principles for determining ARMA model's orders

Model type	Autocorrelation coefficients	Partial autocorrelation coefficients
$AR(p)$	Trailing, index attenuation is slow	p order truncation
$MA(q)$	q order truncation	Trailing, index attenuation is slow
$ARMA(p, q)$	Trailing, index attenuation is slow	Trailing, index attenuation is slow

3. Parameter estimation

   The parameters in $\mathrm{ARMA}(p, q)$ contain $\phi_i \ (i = 1, 2, \ldots, p)$ and $\theta_j \ (j = 1, 2, \ldots, q)$. These parameters are estimated on the basis of maximum likelihood estimation (MLE). Also due to the complexity of $\mathrm{ARMA}(p, q)$, these parameters need to be identified by iteration.

4. Prediction based on ARMA model

   Two prediction methods are available for ARMA model. The first one is multi-step forward prediction according to selected estimation interval. Another one is single-step forward prediction, i.e., using an actual value replaces a predicted value when finishing each prediction.

Energy Consumption Prediction Model of Highway Transportation

We collect data on Unit Energy Consumption from 1999 to 2012 in Inner Mongolia. Unit Energy Consumption in this paper means energy consumption per unit Converted Transport Turnover as mentioned in Introduction. $\mathrm{ARMA}(p, q)$ is used to develop a prediction model of Unit Energy Consumption. EViews 5.0 is the software tool.

1. Stationarity test for Unit Energy Consumption

   $\mathrm{ARMA}(p, q)$ is only suitable for dealing with stationary time series, so we need to firstly test whether the time series (i.e., Unit Energy Consumption from 1999 to 2012) is stationary. By implementing unit root test, the test results are shown in Table 2. It is found from Table 2 that the second-order difference of natural logarithm of Unit Energy Consumption is stationary. Thus, we will develop a prediction model of Unit Energy Consumption based on the second-order difference of natural logarithm of Unit Energy Consumption. Note: Unit Energy Consumption and its first-order difference are not stationary series, so the test results for them have not been shown here.

2. Order determination for Unit Energy Consumption

   The purpose of this step is to identify the values for p and q in the second-order difference of natural logarithm of Unit Energy Consumption. By conducting autocorrelation and partial autocorrelation, the results are shown in Fig. 1. Observing from Fig. 1 and according to Table 1 as well, we find that MA(4) is

Table 2 Stationarity test results for the second-order difference of natural logarithm of Unit Energy Consumption

		t-Statistic	Prob.*
Augmented Dickey-Fuller test statistic		−5.16	0.0024
Test critical values	1% level	−4.20	
	5% level	−3.18	
	10% level	−2.72	

the right model for the second-order difference of natural logarithm of Unit Energy Consumption.

3. Parameter estimation for Unit Energy Consumption
   By conducting MLE, the parameters in MA(4) are estimated for the second-order difference of natural logarithm of Unit Energy Consumption in Table 3.

4. Model diagnosis for Unit Energy Consumption
   After the parameters in MA(4) are estimated, we need to carry out two types of tests, i.e., model significance test and parametric significance test. For the second-order difference of natural logarithm of Unit Energy Consumption, the tests results are shown in Fig. 2. According to Table 3 and Fig. 2, the model based on parameters in Step (3) passes these two types of tests.

5. Unit Energy Consumption prediction model
   According to the estimated parameters in Step (3), we establish the prediction model for the second-order difference of natural logarithm of Unit Energy Consumption, it is as follows:

$$\text{dlec2}_t = -0.0259 + \varepsilon_t - 1.2278\varepsilon_{t-1} + 0.2761\varepsilon_{t-4} \tag{4}$$

where dlec2_t is the tth year prediction value of the second-order difference of natural logarithm of Unit Energy Consumption; ε_t is the tth year random error of the second-order difference of natural logarithm of Unit Energy Consumption; ε_{t-1} is the $(t-1)$th year random error of the second-order difference of natural logarithm of Unit Energy Consumption; ε_{t-4} is the $(t-4)$th year random error of the second-order difference of natural logarithm of Unit Energy Consumption.

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.495	-0.495	3.7482	0.053
		2	0.200	-0.061	4.4171	0.110
		3	-0.434	-0.478	7.9334	0.047
		4	0.253	-0.280	9.2764	0.055
		5	-0.008	-0.036	9.2777	0.098
		6	-0.041	-0.333	9.3260	0.156
		7	0.029	-0.191	9.3549	0.228
		8	-0.003	-0.039	9.3553	0.313
		9	-0.001	-0.255	9.3553	0.405
		10	0.001	-0.138	9.3554	0.499
		11	-0.000	-0.043	9.3554	0.589

Fig. 1 Order determination for the second-order difference of natural logarithm of Unit Energy Consumption

Table 3 Parameter estimation for the second-order difference of natural logarithm of Unit Energy Consumption

Variable	Coefficient	Std. Error	t-Statistic	Prob.
C	−0.025874	0.016716	−1.547827	0.1561
MA(1)	−1.227828	0.042113	−29.15562	0.0000
MA(4)	0.276050	0.032869	8.398526	0.0000
R-squared	0.678074	Mean dependent var		
Adjusted-squared	0.606535			
S. E. of regression	0.178255			
Sum squared resid	0.285974			
Log likelihood	5.393295			
F-statistic	9.478358			
Prob (F-statistic)	0/006094			
Inverted MA roots	0.94 + 0.14i	0.94 − 0.14i	−0.32 + 0.45i	−0.32 − 0.45i

Autocorrelation	Partial Correlation		AC	PAC	Q-Stat	Prob
		1	-0.132	-0.132	0.2677	
		2	-0.081	-0.100	0.3775	
		3	-0.409	-0.446	3.5005	0.061
		4	0.016	-0.173	3.5057	0.173
		5	0.069	-0.089	3.6204	0.305
		6	-0.034	-0.322	3.6521	0.455
		7	0.120	-0.042	4.1383	0.530
		8	-0.031	-0.093	4.1777	0.653
		9	-0.025	-0.237	4.2131	0.755
		10	0.001	-0.025	4.2132	0.837
		11	0.006	-0.075	4.2194	0.896
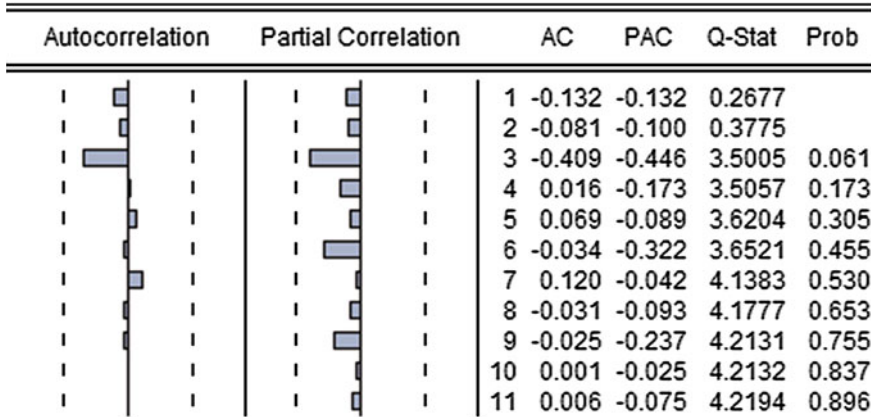
Fig. 2 Residual test for second-order difference of natural logarithm of Unit Energy Consumption

Table 4 Prediction results for the second-order difference of natural logarithm of Unit Energy Consumption in 2013–2020

Year	The second-order difference of natural logarithm of Unit Energy Consumption	Year	The second-order difference of natural logarithm of Unit Energy Consumption
2001	−0.028813	2011	0.004862
2002	0.015435	2012	−0.068996
2003	0.063965	2013	−0.139973
2004	0.073127	2014	−0.030933
2005	0.127569	2015	−0.034149
2006	−0.53064	2016	−0.001822
2007	−0.063136	2017	−0.025874
2008	−0.095627	2018	−0.025874
2009	0.4438	2019	−0.025874
2010	0.009568	2020	−0.025874

Unit Energy Consumption Prediction

The main purpose for modeling Unit Energy Consumption is to forecast energy consumption in future years. Thus, we forecast the Unit Energy Consumption in 2013–2020 using the established model [i.e., Expression (4)].

Due to the dependent variable in Expression (4) is the second-order difference of natural logarithm of Unit Energy Consumption, we need to firstly forecast its second-order difference before forecasting Unit Energy Consumption.

Using the Expression (4), we obtain the prediction values for the second-order difference of natural logarithm of Unit Energy Consumption in 2013–2020, shown in Table 4.

Table 5 Prediction values for Unit Energy Consumption in 2013–2020

Year	UEC (1 kg of SEC/100 converted ton-km)	NUEC	FUEC	SUEC	The prediction values for UEC	The prediction values for NUEC	The prediction values for FUEC	The prediction values for SUEC
1999	5.10	0.7076			5.10	0.7076		
2000	5.07	0.7051	-0.0025		5.07	0.7051	-0.0025	
2001	4.96	0.6958	-0.0093	-0.0068	4.72	0.6738	-0.0313	-0.0288
2002	5.16	0.7128	0.0170	0.0264	5.03	0.7018	0.0061	0.0154
2003	5.24	0.7195	0.0067	-0.0104	6.22	0.7938	0.0810	0.0640
2004	4.78	0.6792	-0.0402	-0.0469	6.30	0.7993	0.0798	0.0731
2005	15.14	1.1801	0.5009	0.5411	5.84	0.7666	0.0873	0.1276
2006	14.59	1.1640	-0.0161	-0.5170	14.14	1.1504	-0.0297	-0.5306
2007	13.02	1.1146	-0.0494	-0.0333	12.16	1.0848	-0.0792	-0.0631
2008	4.79	0.6800	-0.4346	-0.3852	9.32	0.9696	-0.1450	-0.0956
2009	4.61	0.6634	-0.0166	0.4180	4.89	0.6892	0.0092	0.4438
2010	4.35	0.6381	-0.0253	-0.0088	4.53	0.6564	-0.0070	0.0096
2011	3.87	0.5876	-0.0505	-0.0251	4.15	0.6176	-0.0205	0.0049
2012	3.59	0.5553	-0.0323	0.0181	2.94	0.4682	-0.1195	-0.0690
2013					2.42	0.3830	-0.1723	-0.1400
2014					1.51	0.1798	-0.2032	-0.0309
2015					0.88	-0.0576	-0.2374	-0.0341
2016					0.50	-0.2968	-0.2392	-0.0018
2017					0.27	-0.5619	-0.2651	-0.0259

(continued)

Table 5 (continued)

Year	UEC (1 kg of SEC/100 converted ton-km)	NUEC	FUEC	SUEC	The prediction values for UEC	The prediction values for NUEC	The prediction values for FUEC	The prediction values for SUEC
2018					0.14	−0.8528	−0.2910	−0.0259
2019					0.07	−1.1697	−0.3168	−0.0259
2020					0.03	−1.5124	−0.3427	−0.0259

Note
UEC means Unit Energy Consumption
NUEC means the natural logarithm of Unit Energy Consumption
FUEC means the first-order difference of natural logarithm of Unit Energy Consumption
SUEC means the second-order difference of natural logarithm of Unit Energy Consumption
SEC is standard coal equivalent
Converted Transport Turnover (converted ton-km) refers to the sum of freight turnover (ton-km) and the converted freight turnover from passenger turnover (passenger-km)

After we get the prediction values for the second-order difference, we need to convert them into original Unit Energy Consumptions. By difference calculating, the prediction values for Unit Energy Consumption are obtained (Table 5). The predicted Unit Energy Consumption of highway transportation in 2020 in Inner Mongolia is 0.03 kg SEC per 100 converted ton-km.

Conclusions

Based on the data of Unit Energy Consumption of highway transportation in 1999–2012 in Inner Mongolia, a prediction model for Unit Energy Consumption of highway transportation is proposed using $ARMA(p, q)$ model. And the Unit Energy Consumptions in 2013–2020 are forecasted in Inner Mongolia.

From the prediction values for Unit Energy Consumption, we can infer that energy consumptions of highway transportation in Inner Mongolia gradually decline. For the historical energy consumptions (i.e., Unit Energy Consumption in 1999–2012), they also show a declining trend since 2006. Therefore, the prediction values for Unit Energy Consumption in 2013–2020 conform to the historical pattern of energy consumptions.

References

1. Inner Mongolia Autonomous Region Bureau of Statistics. 2010. *Inner Mongolia Statistical Yearbook 2010*. Beijing: China Statistics Press.
2. Inner Mongolia Autonomous Region Bureau of Statistics. 2011. *Inner Mongolia Statistical Yearbook 2011*. Beijing: China Statistics Press.
3. Inner Mongolia Autonomous Region Bureau of Statistics. 2012. *Inner Mongolia Statistical Yearbook 2012*. Beijing: China Statistics Press.
4. Inner Mongolia Autonomous Region Bureau of Statistics. 2013. *Inner Mongolia Statistical Yearbook 2013*. Beijing: China Statistics Press.
5. Inner Mongolia Autonomous Region Bureau of Statistics. 2014. *Inner Mongolia Statistical Yearbook 2014*. Beijing: China Statistics Press.
6. Zhu, Li, Wang Yanxin, and Fan Qibin. 2014. MODWT-ARMA model for time series prediction. *Applied Mathematical Modelling* 38 (5–6): 1859–1865.
7. Zhu, Ke, and Wai Keung Li. 2015. A bootstrapped spectral test for adequacy in weak ARMA models. *Journal of Econometrics* 187 (1): 113–130.
8. Boularouk, Y., and K. Djeddour. 2015. New approximation for ARMA parameters estimate. *Mathematics and Computers in Simulation* 118: 116–122.
9. Cheng, Jing, Dingcheng Zheng, and Jiquan Wu. 2010. Forecasting of Guangdong energy demand based on time-series ARMA model. *Energy Engineering* (1): 1–5.
10. Chu, Chih-Yuan, and Pablo L. Durango-Cohen. 2007. Estimation of infrastructure performance models using state-space specifications of time series models. *Transportation Research Part C* 15 (1): 17–32.

Revenue Model for the Inter-City Railway System Based on the Stop Stations and Graded Ticket Fares

Xichun Chen and Xiaoting Zhao

Abstract With respect to the maximum revenue for the inter-city railway passenger transportation channel system with different grades of parallel trains, this paper has studied the matching relations between the system revenue and the passenger flow demand under confirmed demand conditions, analyzed the behavior selection process of passengers for trains with different speed grades within the channel, confirmed the passenger train flow transformation equation based on Logit sharing rate model as well as the collaborative optimization of stop stations and graded ticket fares, established the maximum revenue model of the system, and designed the hybrid particle swarm harmony search algorithm to solve the model. Besides, the new solution comparative law for the algorithm has also been formulated, which can improve the occurring probability of excellent solutions. Finally, verification has been made by taking Zhengzhou-Xi'an Railway Passenger Transportation Channel as an example, which has showed the effectiveness of the model and the algorithm, and the research results have showed that the redistribution of passenger flow realized through the stop stations and graded ticket fare policy can better improve the maximum revenue of the system.

X. Chen (✉)
School of Traffic and Transportation, Lanzhou Jiaotong University,
Lanzhou 730070, China
e-mail: chenxichun@163.com

X. Zhao
School of Bailie Automotive Engineering, Lanzhou City University,
Lanzhou 730070, China
e-mail: 403382824@qq.com

Introduction

The maximum revenue for the inter-city railway passenger transportation system refers to completing the optimum matching of the existing resource configuration with the passenger flow demand, fully obtaining the revenue of the system through the subdivision of scientific passenger demand, and the graded pricing of train seat on the basis of optimizing the railway operation cost.

Along with the development of high-speed railway as well as the establishment of China Railways Corporation, the revenue management concept extensively applied in the air transportation has been applied in the railway passenger transportation. Shan [1] has studied the aviation reference model for the railway revenue management and established the revenue management model for railway passenger transportation by taking the ticket fare distribution strategy and the price adjustment method as the core. Shi et al. [2] have combined the dynamic ticket fare study for air transportation and the ticket fare management practices of railway transportation market in our country, studied the optimum strategy and practical problems of dynamic railway passenger ticket price of our country, and provided the solution recursion formula for the optimum strategy and practical strategy. The system revenue of the inter-city railway is not only related to the ticket fare, but also closely linked to the utilization rate of train seats. Bao [3] has established the control model and solution algorithm of train seats, and through controlling the train seats, the objective of improving the overall railway revenue can be obtained. The control of train seats and the grading of ticket fare will inevitably affect the selection behavior of passengers and cause the transformation of passenger flow and the optimum matching as well as the maximization of inter-city railway system. Revenue can be realized through the collaborative optimization of the passenger's travel behavior, the control of train seat, and the grading of ticket fare.

Travel Behavior Selection Process for Passengers

The multiple levels of railway travel demand have made the railway have to operate various trains of different service characteristics, so as to meet the demand of passengers, and the train selection of passengers will also influence the compilation of operation schemes.

Based on the aforementioned interactive process, a closed inter-city railway passenger transportation channel has been established, as shown in Fig. 1, so as to analyze the travel selection process of passengers.

Fig. 1 Stop schedule plan of inter-city train

As shown in Fig. 1, there are two grades of passenger trains for the inter-city railway, and all the starting stations as well as the terminal stations thereof refer to A_1 station and A_6 station, G_1 indicates the high-speed passenger train, and the stop station refers to A_3 station and A_5 station, and its targeted passenger type is the high-speed passenger, G_2 indicates the ordinary-speed passenger train, and the stop station refers to A_2 station, A_3 station and A_4 station, and its targeted passenger type is ordinary-speed passenger.

When selecting trains, firstly, the passenger will consider whether there is stop station at the travel destination, and under the condition that the ticket fare is sufficient without consideration of transfer, if the passenger of high-speed passenger flow wants to go to A_4 station from A_1 station and finds that G_1 train has no stop station at the travel destination-A_4 station, the passenger will select G_2 train. After that, the passenger will select the appropriate seat as per the detailed seat pricing and its own traveling utility. If the passenger of high-speed passenger wants to go to A_3 station from A_1 station and finds that G_1 train and G_2 train all have stop stations at A_3 station, the passenger will preferentially select G_1 train, and after that the passenger will select the appropriate seat as per the detailed seat pricing and its own traveling utility. In a similar way, the selection process for the ordinary-speed passenger flow is the same as above.

During the selection process of the aforementioned passenger behavior, different stop station schemes and seat prices will cause the mutual transformation of initial passenger flow for trains and form the dynamic nature of passenger flow; moreover, the dynamic nature of the passenger flow will also exert the feedback effect on the stop station schemes and the seat price of trains and further realize the linkage between the passenger flow demand and the train operation.

Model of Sharing Ratio

The graded pricing of train seat will cause the transfer of various seats for the inter-city train passenger flow and formed the passenger flow sharing ratio of each seat. In the model of optional passenger flow sharing ratio, *Logit* model [4] is the most widely used one, which is based on the random utility theory, to describe the traffic travel behaviors of each traffic participant or group, and comprehensively considers the various factors influencing the passenger's trips.

Assuming that there are a series of optional train seat assembly $Q = \{Q_K | K = 1, 2, \ldots, m\}$, of which, m refers to the total category of seat, and for the passengers, there will be one corresponding traveling utility U^K for each selection of K train seat. Assuming that U^K is composed by two parts, the first part is V^K, which is the measurable part of the system, and the second part is the random error ε, which can reflect the passenger preferences and the observation error, so the utility function for the passengers to select K train seat is the following:

$$U^K = V^K + \varepsilon = \lambda\left(E^K + T^K F^K\right) + \omega C^K \tag{1}$$

In the formula, U^K refers to the generalized cost when the passenger selects K train seat, E^K refers to the ticket fare of K train seat, T^K refers to the travel time of K train seat, and C^K refers to the quantitative description on the comfort of K train seat. In this paper, the quantitative description [5] on the comfort of passengers for taking K train seat is based on $C^K = 0.08E^K$; λ and ω are the weighting coefficients, which can be obtained through the maximum likelihood method [6]; F^K refers to the value of travel time for passengers who take K train seat.

The general type of *Logit* model is as follows:

$$P^K = \frac{\exp\left(-U^K\right)}{\sum_{K=1}^{m} \exp(-U^K)} \tag{2}$$

P^K refers to the passenger flow sharing rate of K seat.

The models are perfected, to eliminate the serious extension of result difference incurred by the exponential growth. The generalized cost function is adopted for the improved model to replace the utility function and obtains the new railway train seat selection model as shown below:

$$P^K = \frac{\exp(-U^K/\overline{U})}{\sum_{K=1}^{m} \exp(-U^K/\overline{U})} \tag{3}$$

In the formula, $\overline{U}$ is the average generalized cost value for m kinds of railway train seats.

$$\overline{U} = \left(\sum_{K=1}^{m} U^K\right)/m \tag{4}$$

Model

Research Scope

The inter-city railway studied in the paper refers to the inter-city railway passenger transportation channel which is jointly constituted by the high-speed railway line and the existing ordinary railway lines, and under the existing operation mode, the passenger number considering the transfer between different grades of trains cannot be strong enough to affect the ticket fare and passenger flow adjustment, so this will not be considered during the study.

Variable Definition

Assuming that the inter-city railway passenger transportation channel has n stations, the starting station is 1, the terminal station is n, the station number from the starting station to the terminal station, respectively, includes $1, 2, 3, \ldots, n$, and i and j are used to mark the stations. The train grade assembly is $S = \{S_1, S_2\}$, of which S_1 is the high-speed train, S_2 is the ordinary-speed train, and $G \in S$. The section assembly is $L = \{l_{ij} | i, j = 1, 2, \ldots, n\}$. The travel time of the section when the passenger selects G train is t_{ij}^G, the standard ticket fare is E_{ij}^G, and the seating capacity is B^G. Assuming that each grade of train can be divided into three seats, the price up-regulation strategy for the upper grade seat is α^G and the price down-regulation strategy for the lower grade seat is β^G. In order to avoid the loss of passenger flow, α^G and β^G will be restricted, and the initial passenger flow for taking G train during l_{ij} section is d_{ij}^G. Due to the graded pricing of train seat as well as the optimization of stop station schemes, which will cause the transformation of passenger flow, the actual passenger flow after the transformation is $\overline{d_{ij}^G}$.

Passenger Flow Transformation Equation

Assuming that θ_i^G indicates whether G train has stop station at i station, and if any, $\theta_i^G = 1$, otherwise $\theta_i^G = 0$; therefore, it can be seen that, θ_{ij}^G indicates whether G train has stop stations at the two ends of l_{ij} section, and if any, $\theta_{ij}^G = 1$, otherwise $\theta_{ij}^G = 0$. The initial passenger flow OD is given, and the passenger flow OD matrix is $D = (d_{ij})_{n \times n}$. The passengers can select the corresponding types of train as their own passenger flow type.

Accordingly, the passenger flow transformation equation based on the stop station scheme and the graded pricing of seats can be listed. Assuming that the initial passenger flow sharing rates of G_1 and G_2 trains are P^1 and P^2, respectively, the passenger flow sharing rates after the passenger flow transformation are, respectively, $P_{\alpha}^1, P^{1'}$ and P_{β}^1 as well as $P_{\alpha}^2, P^{2'}$ and P_{β}^2, and the transformation equation of passenger flow is the following:

$$\overline{d_{ij}^{G_1}} = \theta_{ij}^{G_1} \left[(d_{ij}^{G_1} + d_{ij}^{G_2})(P_{\alpha}^1 + P^{1'} + P_{\beta}^1) + (1 - \theta_{ij}^{G_2})(d_{ij}^{G_1} + d_{ij}^{G_2})(P_{\alpha}^2 + P^{2'} + P_{\beta}^2) \right]$$

(5)

$$\overline{d_{ij}^{G_2}} = \theta_{ij}^{G_2} \left[(d_{ij}^{G_1} + d_{ij}^{G_2})(P_{\alpha}^2 + P^{2'} + P_{\beta}^2) + (1 - \theta_{ij}^{G_1})(d_{ij}^{G_1} + d_{ij}^{G_2})(P_{\alpha}^1 + P^{1'} + P_{\beta}^1) \right]$$

(6)

This equation indicates the changes of actual passenger flow relative to the initial passenger flow when the stop station modes and the passenger flow sharing rate are changed, and the dynamic nature of the equation indicates the constantly and newly generated actual passenger flow during the calculation. Then, the actual passenger flow $\overline{d_{ij}^G}$ shall be taken as the dependent variable, α^G, β^G, and θ_{ij}^G shall be taken as the independent variables. And the calculation shall be made as the above equation, so as to generate the corresponding $\overline{d_{ij}^G}$ assembly.

Objective Function

This paper has summarized the maximum revenue for the inter-city railway passenger transportation channel system as the maximum ticket fare revenue for the inter-city railway as well as the minimum operation numbers of train.

The ticket fare revenue of G train is:

$$M^G = \sum_{i=1}^{n} \sum_{j=i}^{n} E_{ij}^G \left[(1 + \alpha^G)\overline{d_{ij}^G} P_\alpha^G + \overline{d_{ij}^G} P^{G'} + (1 - \beta^G)\overline{d_{ij}^G} P_\beta^G \right] \tag{7}$$

The operation number of G train is the shortest.

Using the maximum section passenger flow to solve the operation number of G train, the objective function is as shown below:

$$\gamma^G = \left\lceil \max_{i<j} \left\{ \sum_{i=1}^{n} \sum_{j=i+1}^{n} \alpha_i^G \cdot \alpha_j^G \cdot \overline{d_{ij}^G} \right\} / B^G \right\rceil \tag{8}$$

In the formula, γ^G indicates the operation number of G train

Therefore, the maximum revenue function of the system is the following:

$$M = \sum_{G=1}^{2} M^G / \gamma^G \tag{9}$$

Constraint Conditions

Limitation on the Transportation Capacity of the Section

Assuming that the upper limit for the traffic capacity of l_{ij} section is $L_{ij}^{\max}$, then the frequency of trains within this section shall not surpass the maximum traffic capacity of the section with the minimum traffic capacity

$$\sum_{G=1}^{2} \gamma^{G} \leq \min \sum_{i=1}^{n} \sum_{j=i}^{n} L_{ij}^{\max} \tag{10}$$

Limitation on the Traffic Capacity of the Station

Assuming that the upper limit for the traffic capacity of the station is $N_i^{\max}$, then the traffic capacity of the station shall not surpass the maximum traffic capacity of the station with the minimum traffic capacity

$$\sum_{K \in T} \theta_i^{G} \gamma^{G} \leq \min \sum_{i=1}^{n} N_i^{\max} \tag{11}$$

Limitation on the Ticket Fare Adjustment Policy

To prevent the passenger flow loss incurred by the adjustment of ticket fare, after the up-regulation, the ticket fare for the high-grade seat shall not be higher than the difference value for the converted air fare within the same section and the values of travel time within the train section.

$$\alpha^{G} \leq \min \sum_{i=1}^{n} \sum_{j=i}^{n} \left[(E_{ij}^{H} - F * \Delta T_{ij}^{H}) / E_{ij}^{G} \right] - 1 \tag{12}$$

In the formula, E_{ij}^{H} refers to the air fare or the converted air fare within the section of $[i, j]$, and ΔT_{ij}^{H} refers to the difference value between the air travel time within the section of $[i, j]$ and the travel time of G train.

After the down-regulation, the ticket fare for the low-grade seat shall not be higher than the difference value between the ticket fare of long-distance bus within the same section and the value of travel time within the train section.

$$\beta^{G} \geq 1 - \min \sum_{i=1}^{n} \sum_{j=i}^{n} \left[(E_{ij}^{B} - F * \Delta T_{ij}^{B}) / E_{ij}^{G} \right] \tag{13}$$

In the formula, E_{ij}^{B} refers to the long-distance bus ticket fare within the section of $[i, j]$, and ΔT_{ij}^{B} refers to the difference value between the travel times of long-distance bus and G train within the section of $[i, j]$.

Limitation on the Parking Coverage

To ensure that the passengers can take the train at each station, restrictions are set to ensure that at least one grade of train can be stopped at each station.

$$\sum_{G=1}^{2} \theta_i^G \geq 1 \tag{14}$$

Limitation on the Maximum Parking Number of the Station

$$\sum_{i=1}^{n} \theta_i^G < n \tag{15}$$

Solution Algorithm

The particle swarm harmony search algorithm (hereinafter referred to as *PSO–HS* algorithm) [7] is adopted for the solution of the model, and improvement is made through utilizing the harmony search algorithm's advantages in jumping out of the local optimum as well as the particle swarm algorithm's features of fast speed and directionality. And the algorithm process is as follows:

Step 1 Confirm the basic parameters of *PSO–HS* algorithm:

The size of harmony memory *HMS*; the number of variable *N*; the solution space of each variable *T* dimension; the harmony memory reservation probability *HMCR*, the evaluation range of tone control probability *PAR*(*PAR*_min, *PAR*_max), the evaluation range of tone fine-tuning disturbance bandwidth *bw*(*bw*_min, *bw*_max); the maximum number of iterations *NI*; the numbers of new solutions generated for each updating of algorithm M_n; the parameters required for the particle swarm algorithm: inertia weight *w*, acceleration constant c_1, c_2, and the maximum number of iterations *MAXNum*; the number of particle; and the size of tone library shall be the same, i.e., *HMS*.

Step 2 Construction of initial particle:

The initial solution *X* of the question refers to the solution vector constituted by the decision variable $x_i(i = 1, 2, \ldots, t, \ldots, T)$, the multi-dimensional (*T* dimension) initial particle is constructed as per the number of the variables, and each dimension will represent the evaluation of each variable of the solution. And as shown in Fig. 2, the part of particle will, respectively, represent the stop station, the ticket fare for high-speed railway seat and the corresponding passenger flow, and the ticket fare for ordinary-speed railway seat and the corresponding passenger flow.

Step 3 Calculate the adaptive value of each particle—the objective function $f(X)$, i.e., the railway enterprise revenue *M*

Fig. 2 Construction of the
initial particle
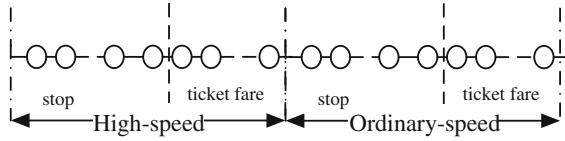
Step 4 Particles updating formula within the particle swarm algorithm

$$x_{t+1} = x_t + v_{t+1}$$
$$v_{t+1} = wv_t + c_1 * rand(pbest - x_t) + c_2 * rand(gbest - x_t)$$

In the formula, *pbest* and *gbest* refer to the local optimum position and the global optimum position of the particle, respectively.

Step 5 The generation of new harmony search algorithm solution

Calculate the former *HMS* optimal solutions within the particle swarm, take them as the initial solution of *HS* algorithm, and it shall be stressed that the result obtained by *PSO* is the initial solution of harmony search algorithm, but the initial solution shall be the former *HMS* solution that are most optimal in the particle swarm algorithm

(1) Calculate the best particle *xbest* and the worst particle *xworst* within each generation.
(2) As per the current iteration conditions, calculate its self-adaption, i.e., *bw* and *PAR*.

The *sigmoid* function is adopted to change the inertia weight, for it possesses smoother top and bottom than the cosine function, and can better balance the liner and nonlinear behaviors. *sigmiod* function can describe the variation tendency of the inertia factor *w* along with the algorithm evolution, and the evaluation interval of *sigmoid* function is [−9.903438, 9.903438].

Decreasing formula for inertia weight is $w = [(w_{max} - w_{min})/1 + e^{(20G/G_{max}-10)}] + w_{min}$

$$PAR = PAR_{min} + (PAR_{max} - PAR_{min}) * ni/NI$$
$$bw = bw_{max} \exp(\log bw_{min}/bw_{max}) * ni/NI$$

(3) Using *bw* and *PAR* to obtain the current solution of harmony search algorithm, i.e., x_{1i}

Step 6 Memory bank updating

(1) Randomly select one solution element from each dimension ranging from the first dimension to *b* dimension, during which a higher rate of convergence is required in the face of complicated actual problems, which is the weakness of harmony search algorithm, and through introducing the combination of particle

swarm algorithm and harmony search algorithm again, the rate of convergence for the solution can be increased, and the combination formula is:

$$x_{t+1} = x_t + rand(pbest - x_t)$$

Use the formula to obtain the current solution, i.e., x_{2i}.

(2) Compare $f(x_{1i})$, $f(x_{2i})$, and $f(xworst)$, and if $f(x_{2i}) \geq f(x_{1i})$ & $f(x_{2i}) \geq f(xworst)$, x_{2i} shall replace $xworst$, and if $f(x_{1i}) \geq f(x_{2i})$ & $f(x_{1i}) \geq f(xworst)$, x_{1i} shall replace $xworst$.

Step 7 Iteration termination

Judge whether the number of iterations has reached to the maximum value NI, and if yes, stop the iteration and output the optimal solution; otherwise, repeat the steps (5) and (6).

Case Analysis

This paper has made verification by taking Zhengzhou-Xi'an Inter-city Railway Passenger Transportation Channel as an example, and two transportation modes are selected, i.e., the high-speed train and the ordinary-speed train. And the calibration of relevant parameters is as follows:

1. The traffic capacity of the section: 40 pairs of high-speed trains and 25 pairs of ordinary-speed trains passed every day and night, 60 pairs of trains with the traffic capacity of the starting station and terminal station as well as 45 pairs of trains with stop stations within the section are taken in the case.
2. The values of travel time for passengers: As per the 3780 copies of investigation sample data considering the travel selection of passengers with different incomes in Zhengzhou-Xi'an Transportation Channel, the values for the travel time of high-speed train and ordinary-speed train are, respectively, RMB 30/h and RMB 18/h according to the calculations.
3. The evaluation of benchmark ticket fare: the second-grade seat for the high-speed train on the train schedule of Zhengzhou-Xi'an Passenger Transportation Channel is taken as the benchmark price, and the hard sleeper of ordinary-speed train is taken as the benchmark ticket fare.
4. The basic parameters for the operation of train are as shown in Table 1.
5. Utility function parameter calibration.

As per the 3780 copies of investigation sample data considering the travel selection of passengers with different incomes in Zhengzhou-Xi'an Transportation Channel, the maximum likelihood method is adopted to calibrate the model parameter, and *Matlab* software is applied to obtain the maximum optimal parameter of the log-likelihood function, i.e., the parameter for the travel utility

Table 1 Basic parameters of train operation

Type	High-speed train	Ordinary-speed train
Seating capacity (person)	610	1114
Average speed of the train (km/h)	250	72

function of each income group in Zhengzhou-Xi'an Passenger Transportation Channel, and the solution includes $\lambda = 0.019$ and $\omega = 0.078$.

As per the passenger flow statistic conditions of the current Xi'an Railway Bureau and Zhengzhou Railway Bureau, the passenger flow data of 6 stations including Zhengzhou, Gongyi, Luoyang, Sanmenxia, Weinan, and Xi'an are selected, and the obtained initial matrix is as follows:

$$d_{ij}^{G_1} = \begin{bmatrix} 0 & 110 & 1870 & 1000 & 980 & 4250 \\ & 0 & 480 & 280 & 200 & 690 \\ & & 0 & 870 & 720 & 2750 \\ & & & 0 & 500 & 1660 \\ & & & & 0 & 850 \\ & & & & & 0 \end{bmatrix}$$

$$d_{ij}^{G_2} = \begin{bmatrix} 0 & 580 & 1540 & 860 & 520 & 2850 \\ & 0 & 850 & 350 & 100 & 900 \\ & & 0 & 910 & 180 & 1080 \\ & & & 0 & 220 & 750 \\ & & & & 0 & 540 \\ & & & & & 0 \end{bmatrix}$$

The inter-station travel time matrix for the high-speed train/ordinary-speed train of Zhengzhou-Xi'an Railway Passenger Transportation Channel

$$T^{G^1/G^2} = \begin{bmatrix} 0/0 & 20/56 & 38/94 & 71/199 & 120/329 & 147/379 \\ & 0/0 & 18/38 & 51/143 & 100/273 & 127/323 \\ & & 0/0 & 30/105 & 86/235 & 108/285 \\ & & & 0/0 & 54/130 & 76/180 \\ & & & & 0/0 & 20/50 \\ & & & & & 0/0 \end{bmatrix}$$

As per the aforementioned particle swarm harmony search algorithm, *matlab* programming is adopted for the solution, and the calculation results are as follows: the stop stations for the high-speed railway include Zhengzhou, Luoyang, Sanmenxia, Weinan, and Xi'an stations, the ticket fare rate for the first-grade seat is 0.511, and the number of opening line is 10 pairs. The stop stations for the ordinary-speed train include Zhengzhou, Gongyi, Luoyang, Sanmenxia, and Xi'an

stations, the ticket fare rates for the soft sleeper and hard seat are, respectively, 0.625 and 0.375, the number of opening line is 8 pairs, and the daily revenue of the railway passenger transportation channel is RMB 1,931,723. After the transformation, the passenger flow matrix for the first and second grades of seat of the high-speed railway as well as the soft sleeper, hard sleeper, and hard seat of the ordinary-speed railways are as follows, and since the high-speed train and the ordinary-speed train are all set with nonstop station restrictions, the passenger flow between Gongyi and Weinan stations has been optimized.

$$
\overline{d_{ij}^{G_1}} = \begin{bmatrix} 0 & 0/0 & 455/1552 & 259/868 & 345/1155 & 966/3235 \\ & 0 & 0/0 & 0/0 & 0/0 & 0/0 \\ & & 0 & 262/876 & 207/693 & 524/1756 \\ & & & 0 & 166/554 & 321/1073 \\ & & & & 0 & 320/1070 \\ & & & & & 0 \end{bmatrix}
$$

$$
\overline{d_{ij}^{G_2}} = \begin{bmatrix} 0 & 83/276/331 & 168/561/674 & 88/293/352 & 0/0/0 & 348/1160/1391 \\ & 0 & 160/532/638 & 76/252/302 & 0/0/0 & 191/636/763 \\ & & 0 & 77/257/308 & 0/0/0 & 186/620/744 \\ & & & 0 & 0/0/0 & 122/406/488 \\ & & & & 0 & 0/0/0 \\ & & & & & 0 \end{bmatrix}
$$

When comparing the calculation result with the existing Zhengzhou-Xi'an Passenger Transportation Channel, it can be found that the stop station scheme thereof is consistent, indicating that this model possesses good practicability, and through taking the ticket fare between Zhengzhou and Xi'an stations as an example, the ticket fare for the hard seat and the ticket fare for the soft sleeper of ordinary trains are increased in the new ticket fare strategies, of which the ticket fare for hard seat is RMB 85, which is RMB 13 higher than the existing fare; and the ticket fare for soft sleeper is RMB 228, which is RMB 24 higher than the existing fare; and for the ticket fare of the first-grade seat of high-speed railway, it is RMB 346, which is RMB 23 lower than the existing fare. When comparing these data with the ticket fare for aviation and long-distance bus within the same distance, the fluctuation of the aforementioned ticket fares will not change the selection behavior; meanwhile, due to the adjustment of ticket fare, the demand for the first-grade seat of high-speed train has been increased to some extent, and the demand for sleepers of the ordinary-speed trains have also been largely increased by RMB 253,929 than the additional ticket fare before the adjustment; therefore, this has exerted an obvious effect on the improvement of railway enterprise revenue. Thus, we can largely increase the total system revenue of the inter-city railway passenger transportation channel through adjusting the stop station mode and the graded ticket fare policies, and meanwhile, fully digging the potential revenue space as per the demand features of the passenger flow.

Conclusion

This paper takes the inter-city railway passenger transportation system revenue as the research objective, analyzes the influences of stop station and graded prancing of train seats on the transformation of passenger flow demand, obtains the passenger flow transformation equation, and describes the internal mechanism of the passenger flow transformation by taking the behavior selection process of the passenger as the turning point.

The established revenue model for the inter-city railway system has optimized the matching relations among the stop stations of train, the classification of train seats, and the inter-station passenger flow demand. The solution of the model can be calculated through the hybrid particle swarm harmony search algorithm, and new solution comparative laws has been established for the algorithm, which can improve the occurrence probability of excellent solutions.

This example has showed the effectiveness of the model and algorithm in this paper, and it is a beneficial attempt to use the graded pricing of stop station and train seat to affect the redistribution of passenger flow demand. The next research emphasis is how to confirm the operation scheme through the uncertainty demand of passengers, so as to realize the optimum matching of railway department revenue with the passenger flow demand.

References

1. Shan, X. H. 2012. *Study on railway revenue management model and its application in passenger transportation*, 20–29. Beijing: China Academy of Railway Sciences.
2. Shi, F., G. H. Zheng, and Q. Gu. 2002. Optimal dynamic pricing of railway passenger ticket. *Journal of the China Railway Society* 24 (1): 1–4.
3. Bao, Y. 2013. *The theory and methods for railway seat inventory control*, 113–128. Beijing: Beijing Jiaotong University.
4. Nuzzolo, Agostino, Umberto Crisalli, and Francesca Gangemi. 2000. A behavioural choice model for the evaluation of railway supply and pricing policies. *Transportation Research Part A: Policy and Practice* 34 (5): 395–404.
5. Wu, W. X. 2011. Research on the share rate of passenger flow and transportation organization strategy in railway transportation corridor. *China Railway Science* 32 (2): 126–129.
6. Mi, Jie. 2006. MLE of parameters of location-scale distribution for complete and partially grouped data. *Journal of Statistical Planning and Inference* 136: 3565–3582.
7. Mahdavi, M., M. Fesanghary, E. Damangir. 2007. An improved harmony search algorithm for solving optimization problems. *Applied Mathematics and Computation* 188 (2): 1567–1579.
8. Peng, H. Q., and Y. J. Zhu. 2012. Intercity train operation schemes based on passenger flow dynamic assignment. *Journal of Transportation Systems Engineering and Information Technology* 47 (3): 484–489.

Four-Phase Composite Material of Concrete Meso-Damage Dynamic Load Failure Test

Ji-kun Zhao, Shu-jun Huang and Shui-zhong Tao

Abstract At mesoscopic scale, the research proposed and established that concrete was a four-phase continuous and inhomogeneous composite material model consisting of coarse/fine aggregates, hydration products in cement, pores, and cracks. Under static and dynamic loads, micro-mechanics damage test simulation was based on the four-phase composite material model of concrete. The research's content was concrete uniaxial compression test, three-point bending girder's damage, and fracture under static and dynamic loads. The research found that the location of main crack in the beam shear was near the frailest part. The crack was along the bone and sand slurry's combination, extending to the loading point direction and showing tortuosity. The crack extended surface was uneven, which was temporary deflection of the crack propagation direction due to the heterogeneity of concrete and aggregate distribution in it. The stress–strain curves of static and dynamic loads were similar, and the starting points of fracture initiation were basically identical. The stress wave in the specimen propagated back and forth many times, which led to the specimen fracture. Under the dynamic load, the elastic modulus of concrete was higher than that under the static load, and the numerical simulation results agreed with the experimental results. When the strain rate was less than a critical value, concrete strength increased slowly. When the strain rate was more than the value, concrete strength increased quickly. This research established five-layer, eight-layer, and 10-layer concrete frame structural models, which were simulated at the damage process of frame structure in strong earthquake. The study found that beam and column alternating failure belongs to the mixed collapse mechanism. The weaker the stiffness of foundation was, the earlier

J. Zhao (✉) · S. Huang
College of Engineering, Nanjing Agricultural University, Nanjing 210031, China
e-mail: jikunzhao_2006@163.com

S. Tao
Beijing Municipal Engineering Research Institute, Beijing 100037, China

the failure time of lower floors was and the deeper the structural settlement was. The stronger the stiffness of foundation was, the earlier the failure point generated. The low floor failure of frame structure consumed a part of seismic energy, which reduced the effect of earthquake on upper frame structure partly.

Keywords Concrete · Four-phase composite materials · Elastic-plastic damage · Random defects · Mesh damage elimination

Introduction

At mesoscopic scale, concrete is a multiphase continuous and inhomogeneous composite material consisting of coarse/fine aggregates, hydration products in cement, unhydrated cement particle, pores, and cracks. Therefore, the researcher put forward some numerical simulations to study mesoscopic problems of concrete. Du et al. [1] constructed the random aggregate model by using Monte Carlo method, in which the concrete was considered as a three-phase composite material composed of cement mortar, aggregate, and the interface between the two. Other researchers [2–5] also simulated concrete as a three-phase composite material, but the aggregate shape and deposition method were different. In fact, there are many original micro-cracks or micro-defects randomly distributed in concrete, whose lengths are less than 0.5 mm mostly. The micro-crack has a certain direction and is a result from comprehensive factors. So it needs to increase the simulations of random micro cracks or defects, to accord with actual the concrete material preferably. Concrete is a quasi-brittle material; hence, the Linear Elastic Fracture Mechanics cannot be directly applied to concrete. Research shows that the load-displacement nonlinear showing of concrete is closely related to the development of fracture process zones [6]. The deformation and failure process of concrete under the dynamic load have an important effect on studies closely related to engineering problems, such as concrete crushing. Along with the progress of experimental and test method, a large number of experimental studies [7] have been carried out to discuss the relationship between the elastic modulus or strength of concrete and the deformation rate under the dynamic load. And the strength criterion and constitutive relation of concrete under the dynamic load have been summarized. Those lay the foundation for studying dynamic properties and mechanical models of concrete.

The research proposed that concrete was a four-phase composite material micro-mechanics model consisting of aggregate, mortar, interface, and random defects. Numerical samples of three-point bending girder composed of the four-phase composite material of concrete were set up. Under static and dynamic loads, damage test on each sample was performed. It also established different layers of frame structural models, which were simulated to analyze collapse test under the seismic load.

Establishment of Concrete Elasto-Plastic Damage Constitutive Model

Realization of the Elastic-Plastic Damage Constitutive Equation of Mesoscopic Element

Damage mechanics model of concrete includes Mazars damage model, Loland damage model, and statistical damage model. On the basis of damage mechanics theory and plastic theory, the constitutive relation can be obtained considering the elastic-plastic damage in strain space [8, 9]. Under the assumption that the total deformation of concrete mesoscopic element in the damaged section consists of elastic deformation and plastic deformation in three-dimensional stress state, and elastic deformation obeys generalized Hook's law, the stress can be written as

$$\sigma = (1 - W)D_0^e : (\varepsilon - \varepsilon^p) \tag{1}$$

where D_0^e is the elastic fourth-order tensor under the undamaged condition, W is the damage variable, and ε is the total deformation tensor, generally divided into two parts, elastic strain tensor ε^e and plastic strain tensor ε^p, that is $\varepsilon = \varepsilon^e + \varepsilon^p$.

In order to apply the deduced elastic-plastic damage constitutive equation to the mesoscale concrete sample, the sample must meet the mesoscopic unit size. In addition, based on Weibull distribution [3], it can assign the value to mesoscopic element material, which can help to describe the nonuniformity of material properties. The probability density function of Weibull distribution is

$$g(u) = \frac{m}{u_0} \left(\frac{u}{u_0} \right)^{m-1} \exp\left[-\left(\frac{u}{u_0} \right)^m \right] \tag{2}$$

where u is the random variable meeting this distribution (such as strength, elastic modulus, and Poisson's ratio), u_0 is the parameter related to mean value of random variables, and m is the homogeneous degree, and it can determine the shape of probability density function of Weibull distribution.

Establishment of Concrete Random Defects

There are many micro-cracks or micro-defects in concrete inner, and those defects may be existed in original materials, or induced by late stress. Under the certain loading condition, those micro-defects will be further expanded. The formation and extension of micro-defects are all directional. The micro-defects' closure and friction in compression cause very complicated mechanical behavior. So the numerical simulation on micro-defects is very important. In three-dimensional space, under the assumption that the center $O(x^0, y^0, z^0)$ location of random

micro-cracks or micro-defects in concrete inner distributed randomly, the concrete internal random micro-cracks or micro-defects meet the following curve function:

$$f_i = \left(x_i - x_l^0\right)^2 / a_i = 1 \quad (i = 1, 2, 3) \tag{3}$$

where a_i is a random variable representing the length of half axis, b_i is the initial length of half axis, and $a_i = b_i r_j$, in which r_j is the uniform random number on (0, 1) interval.

Equation (3) is the formula for the lth random micro-cracks or micro-defects. The mesoscopic unit that satisfies Eq. (3) is the concrete random micro-defect unit. Need to be explained that, the structure of random micro-defects that this paper established was unconnected, different from macroscopic pre-crack. And it weakens the features of original defects, which cause the mechanical properties of material lower than those of the interfacial properties of concrete mesoscale structure. Of course, it needs to be determined according to the test results. And then it establishes concrete four-phase composite material model, the more analysis in the literature [10].

Concrete Failure Test Under Static And Dynamic Loads

Concrete is a quasi-brittle material, which is sensitive to tensile stress. So it is very important to research on concrete tensile fracture. Tensile fracture, which is the extension of crack type I in fracture mechanics, is a mature part of the research on linear fracture mechanics. However, there are still many problems when using classical fracture mechanics to research concrete material, such as concrete is inhomogeneous that does not satisfy the material uniformity hypothesis in fracture mechanics, and there are fracture process zones in concrete fracture process, etc. The specimens used to study the tensile fracture can be divided into direct tensile specimen, three-point (or four-point) bending specimen, and splitting specimen. According to the research of fracture mechanics, one or two initial cracks are precast in the specimen, to ensure that fracture is beginning to extend from the pre-crack tip [11, 12].

Establishment of Numerical Sample of Concrete Beam

Under the condition that two-dimensional concrete beam did not precast cracks, this paper researched on cracks extension and strength characteristics in elastic-plastic damage process of three-point bending concrete beam. The span of sample is l, the span–depth ratio is $l/d = 3$, the span size of sample is 125 mm, and the geometry and stress condition of sample are as shown in Fig. 1.
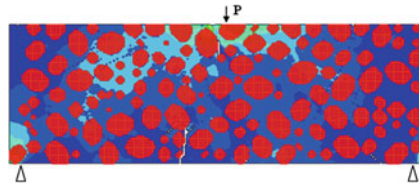
Fig. 1 2D sample of four-phase composite concrete beam

Table 1 Weibull distribution parameters of concrete beam sample

Component	Mean value of elastic modulus (/GPa)	Mean value of strength (/MPa)	Homogeneous degree (m)	Proportion of each component (/%)
Mortar matrix	28.6	175	3	55.82
Interface	15.0	150	1.5	4.12
Aggregate	80.0	500	6	38.88
Random defect	1	0.1	3	1.18

Table 1 shows Weibull distribution parameters of two-dimensional concrete beam sample of four-phase composite material model. Tension–compression ratio of each phase material mesoscopic unit is 1/12. All other parameters are the default values. Generated concrete numerical sample is as shown in Fig. 1.

Concrete Failure Test Under the Static Load

The established two-dimensional concrete numerical sample of four-phase composite material model used displacement loading method of 0.002 mm/step for numerical simulation of elastic-plastic damage, it obtained the displacement loading curve as shown in Fig. 2, and elastic-plastic damage processes were as shown in Fig. 3.

Fig. 2 Displacement-loading curve of concrete beam

(a)   (b)   (c)

      30th step   40th step   70th step
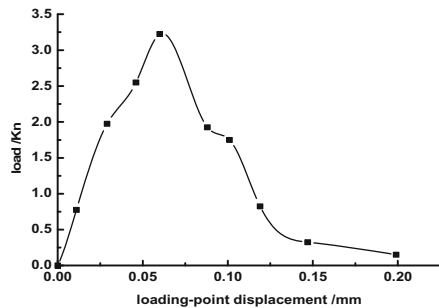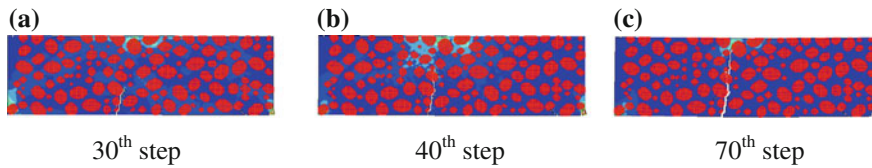
Fig. 3 Fracture process of three-point bending concrete beam of four-phase composite material

As shown in Fig. 2, the three-point bending concrete beam showed a certain degree of elasticity in the macro at the initial stage of loading. With the increase of unit load, some mesoscopic units reached the threshold value of plastic strain damage, appeared damage evolution, and showed as the plastic strengthened stage in Fig. 2. When the load was added to 30 steps, the displacement loading curve showed the softening stage. At this point, the thirtieth step one in Fig. 3 showed a main crack which had been extended to the loading point obviously. This crack was also through the medial axis of the concrete beam sample. With the increase of the load point displacement, concrete beam was softened constantly. The main crack had passed through the whole sample, and the sample had lost the bearing capacity.

As shown in Fig. 3, the tensile crack was distributed in the tension area of bottom surface of concrete beam. The increasing displacement load and the tensile cracks continued to expand and coalesced with each other, forming a main crack. The location of main crack in the two-dimensional concrete beam shear was near the frailest part. The crack was along the bone and sand slurry's combination, and extending to the loading point direction. So the crack propagation path showed the tortuosity. Because of the stress characteristics of two-dimensional concrete beam and the location of random micro-cracks, the main crack had no obvious effect on the whole damage process.

Under the static load, the main crack was generated in the tension area of bottom surface of concrete sample, extending to the loading point direction. In the numerical simulation, the failure phenomena that location of main crack deviated from the symmetric cross section were because of the inhomogeneity of the mesoscale four-phase composite concrete. Compared with the numerical simulation results of concrete sample with precast cracks, the strength of sample with no precast cracks was higher. After the formation of main crack, the fracture characteristics were similar to those of the sample with precast cracks, but the fracture brittleness was more obvious. At the same time, the main crack throughout the whole sample height and the whole concrete beam sample lost the bearing capacity.

Concrete Failure Test Under the Dynamic Load of I

Under the dynamic load, a stress wave changed with time was applied to the loading plate in the upper part of sample. The waveform of applied impact stress
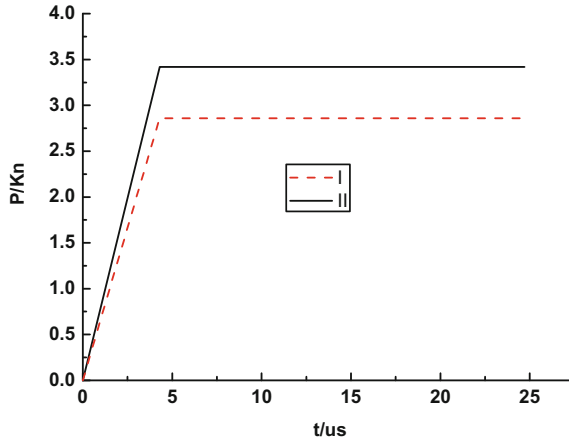
Fig. 4 Dynamic load applied to the upper part of sample
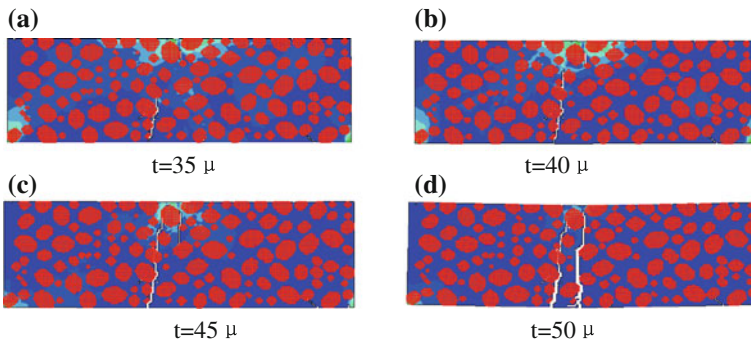
Fig. 5 Fracture process of concrete sample under the dynamic load of I

changed according to two situations (I, II) given in Fig. 4, in order to study the effect of different dynamic stress amplitude on the fracture mode of concrete sample. Under these two loading situations, applied dynamic stress amplitude was 3, 3.5 MPa. Of course, the kinds of loading situations were very difficult to be realized in dynamic impact experiments, such as the Hopkinson bar experiment. But in numerical simulation, the stress wave can be directly applied to the sample, to study more complex stress state.

As shown in Fig. 5, in the situation I, before the stress wave was propagated to the loading plate in the upper part of sample (t = 30 μs), only a few mesoscopic units in the upper part of sample were damaged. When t = 40 μs, due to the multiple reflection and refraction of the stress wave, tensile fracture of many units occurred in the lower part of sample and developed into a few cracks. With the passage of time, stress wave propagated in the upper part of sample, and those

cracks also extended to the upper. Finally, a crack band was formed vertically in the middle of sample. The final formation of fracture mode was similar to that of static load (in Fig. 3). However, under the dynamic load, more cracks were generated in the sample, and their expansion directions were not single (Fig. 5).

Concrete Failure Test Under the Dynamic Load of II

Under the dynamic load of II, the time of unit damage in the sample was greatly advanced, as shown in Fig. 6. When t = 10 μs, cracks were generated near the loading plate in the upper part of sample. The cracks were caused by the compression stress applied directly. Along with the increase of stress, the cracks extended to the lower part of sample, developed into many cracks, and led to the failure of sample.

In summary, under static and dynamic loads, fracture modes of concrete samples were similar basically. What's more, the stress–strain curves of static and dynamic loads were similar, and the starting points of fracture initiation were basically identical. When the applied peak stress was lower, the stress wave in the sample propagated back and forth many times, which led to the sample fracture, and the fracture shape of sample was similar to that of the static load (in Fig. 3). With the increase of the amplitude of stress wave, the fracture mode of sample was also changed. By comparing Figs. 5 and 6, there were some differences between the two fracture modes.

Under the dynamic load, for the stress waves of different peak stresses, the crack initiation and propagation processes were different, and the final fracture modes were different too. And the elastic modulus and Poisson's ratio of concrete changed. The dynamic elastic modulus was higher than that of the static elastic modulus, but the Poisson's ratio was just the opposite. In this numerical simulation, the dynamic and static elastic moduli of five kinds of different concretes were simulated, respectively. Compared with the experimental results of the [13] test, the results showed that the simulation model was reasonable (Fig. 7).

(a) t=15 （elastic modulus）
(b) t=20 （elastic modulus）
(c) t=25 （elastic modulus）
(d) t=30 （elastic modulus）
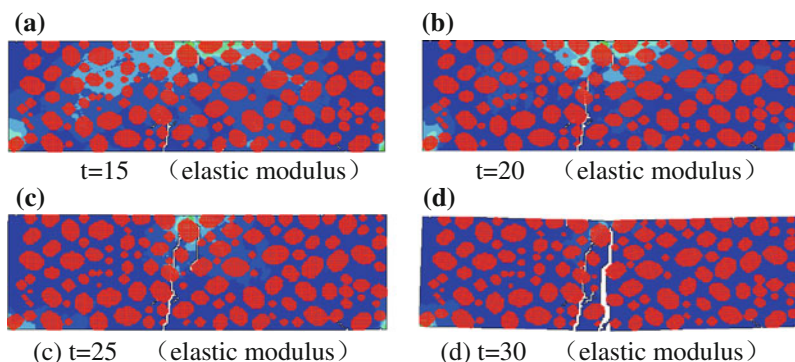
Fig. 6 Fracture process of concrete sample under the dynamic load of II
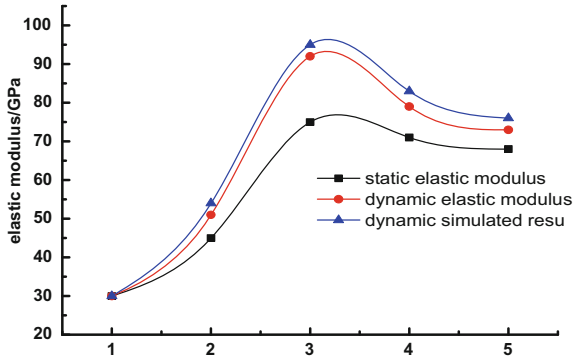
Fig. 7 Dynamic elastic modulus of numerical simulation and test

Damage Test of Concrete Beam Column

This research established five-layer, eight-layer, and 10-layer concrete frame structural three-dimensional models, as shown in Fig. 8. With the simulation of collapse failure of concrete frame structure under earthquake action, it dynamically

5 layers frame structure   8 layers frame structure

10 layers frame structure
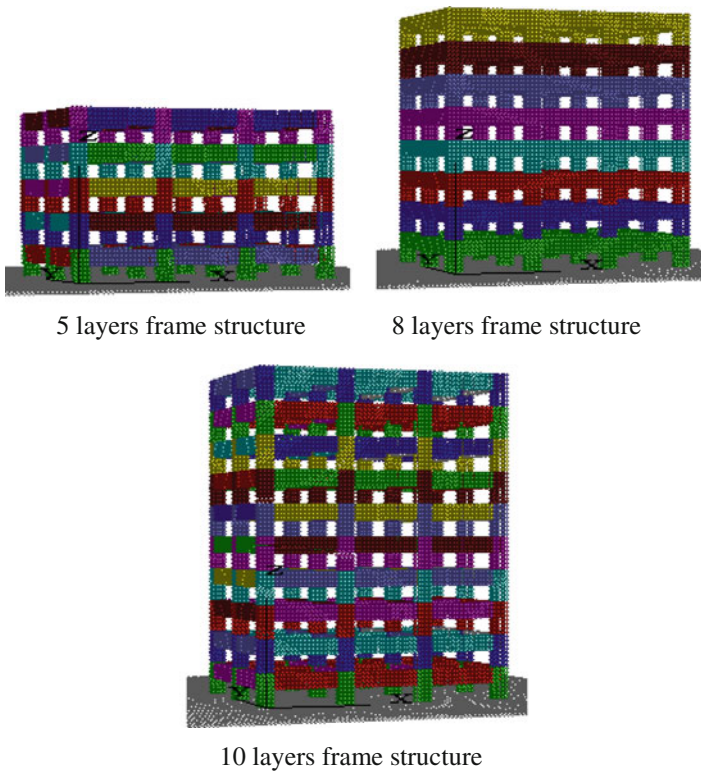
Fig. 8 Frame structural models

displayed that the whole process of seismic response of frame structure from elastic, cracking, yielding to collapse. The influence of structural stiffness and number of layers on the collapse failure of concrete frame structure under the condition of rigid foundation was studied. The top acceleration time history curve of five-story frame structure was as shown in Fig. 9a. The story drift curve of each layer of five-story frame structure was as shown in Fig. 9b. The collapse failure of five-story frame structure was as shown in Fig. 10. The top story drift curves of three kinds of structure were as shown in Fig. 11.

The structure that was in the elastic state before 0.7 s can be obtained from Figs. 8 and 9. With the continuous accumulation of damage, the failure points were firstly appeared at the bottom of beam supported by a column which was in the right side frame structure at the bottom level at 1.25 s (in Fig. 10a). Then, the failure points were appeared at the bottom of column and coalesced with each other quickly, forming cracks at the bottom of all columns at the bottom level of frame structure at 1.44 s. During this time, cracks started at the top of the right side of

Fig. 9 Top acceleration time history curve and story drift curve of five-story frame structure. **a** Top acceleration time history curve. **b** Story drift curve

(a)

5 layers frame structure

(b)

5 layers frame structure

(c)

5 layers frame structure
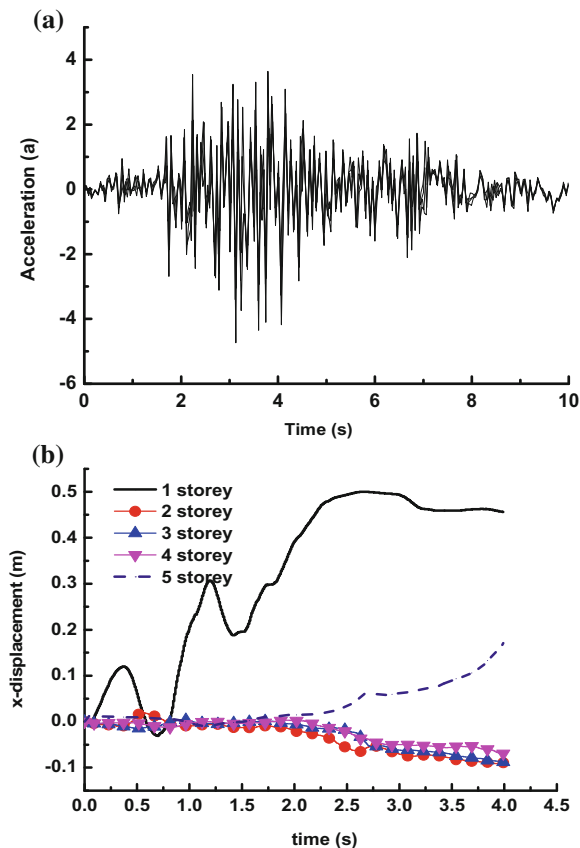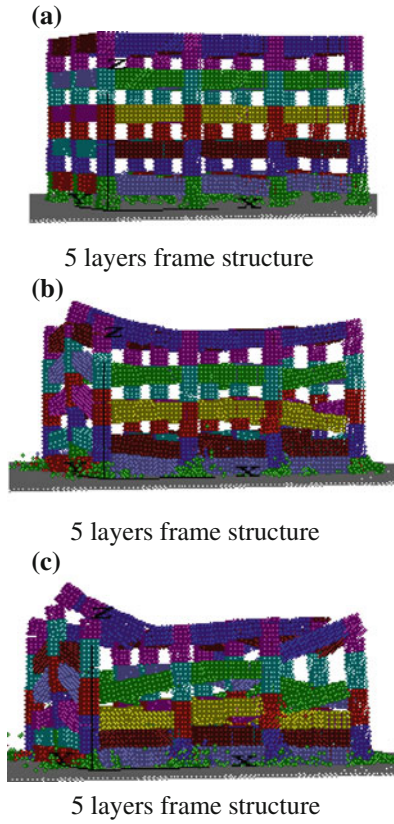
Fig. 10 Collapse process of five-story frame structure
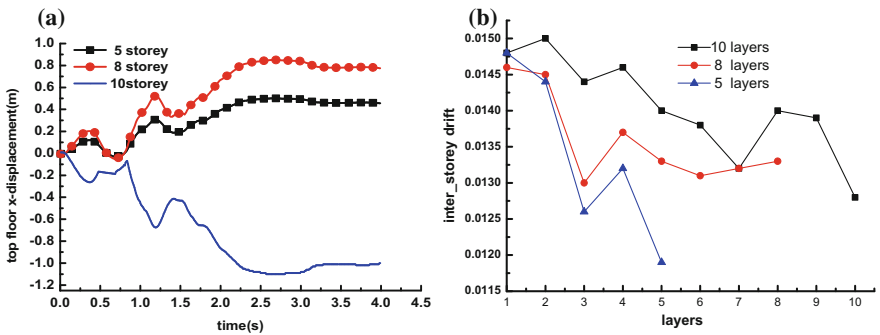
Fig. 11 Story drift curve of frame structure. **a** story drift, **b** inter-story drift

beam at the bottom layer to the third layer down. At the same time, the beam bottom of the third layer also appeared the phenomenon of cracking. At 1.45 s, the second floor and the end region of horizontal beam lost the support and began to collapse. At 1.48 s, the bottom of column of the third layer began to crack. Then, the top and bottom of column of the second floor, the middle section of column of the third floor, the bottom of column of the fourth floor, all appeared cracks and continued to expand. Since 1.6 s, the second floor to fifth floor collapsed successively. At 1.79 s, the frame column staggered and tilted along the bottom crack. At 2.3 s, the beam floor cracks of the second and third floor coalesced with each other, and the floors which lost the support began to collapse. At 3.12 s, the fourth and top floor started to collapse down. As the deformation continued to develop, the frame was collapsed and destroyed completely.

By comparing the top story drift curves of three kinds of frame structures (in Fig. 11), the higher the number of the layer was and the larger the story drift was, the more obvious the tremor under the earthquake was. By comparing the failure processes of eight-layer and ten-layer frame structure, at 1.65 s, columns of the bottom, second and third floors for eight-layer frame structure damaged seriously, and part of the columns began to tilt. But the similar damage phenomenon for ten-layer frame structure was about 0.15 s early. The whole structure lost stability and began to overall collapse at 1.83 s for eight-layer frame structure and at 1.71 s for ten-layer frame structure. With the deformation continued, the first to third floor damaged seriously, and the end region of beam column of upper layers appeared cracks, but the model did not collapse. By comprehensive comparison, beam and column alternating failure belongs to the mixed collapse mechanism. The weaker the stiffness of foundation was, the earlier the failure time of lower floors was and the deeper the structural settlement was. The stronger the stiffness of foundation was, the earlier the failure point generated. The low floor failure of frame structure consumed a part of seismic energy, which reduced the effect of earthquake on upper frame structure partly.

Conclusion

The research proposed and established that concrete was a four-phase composite material model. The failure simulation of three-point bending girder under the dynamic load and shear failure process were studied. The research also analyzed the failure simulation of five-layer, eight-layer, and 10-layer frame structure in strong earthquake. The main conclusions are as follows:

(1) The numerical simulation results find that the crack extended surface is uneven, which is temporary deflection of crack propagation direction due to the heterogeneity of concrete and aggregate distribution in it, instead of the completely destroyed material layer caused by damage with the macro-size. It is very similar to the experimental results.

(2) The dynamic mechanical properties and static mechanical properties of concrete are mainly different from that the dynamic strength of concrete varies with the strain rate. The experiment results prove that when the strain rate is less than a critical value, concrete strength increase slowly. When the strain rate is more than the value, concrete strength increases quickly. In general, under the dynamic load, the sample will produce more cracks, which have different effects on the damage of samples, and it is different from forming a main crack under the static load.

(3) The higher the floor is, the larger the displacement is, the more obvious the tremor under the earthquake is, the shorter the time of collapse is. When the number of layers of frame structure is more than eight, the first to the third layer collapses seriously. And the beam and column alternatively failure. The low floor failure of frame structure consumes a part of seismic energy, which reduces the effect of earthquake on upper frame structure partly.

References

1. Du, X.L., R.J. Tian, and Y.J. Peng. 2007. Application of homogenization on the uniaxial tension test for the numerical concrete. *Journal of Shenyang Jian zhu University (Natural Science)* 23 (5): 742–746.

2. Tang, X.W., and C.H. Zhang 2008. Simulation of meso-fracture for concrete based on the developed random aggregate model. *Journal-Tsinghua University (Science & Technology),* 48 (3): 348–352.

3. Tang, CH. A., and W. CH. Zhu, 2003. Concrete damage and fracture numerical test. *Science Press*.

4. Zhao, J.K., H.Q. Zhang, and S.H.J. Huang. 2016. Seismic response analysis of rural masonry building based on push-over. *Journal of Nanjing Agricultural University* 39 (2): 341–348.

5. Lin, G., J.B. Li, J. Zhao, et al. 2007. Mesoscopic numerical evolution analysis of concrete damage under uniaxial tension and compression. *Journal of Architecture and Civil Engineering* 24 (1): 1–6.

6. Wu, Z.M., W. Dong, and Q. Xu. 2009. Propagation criterion for mixed modeI-IIcrack propagation and numerical simulation of the whole propagation process. *Journal of Hydraulic Engineering* 40 (2): 180–187.

7. Trisha Sain, J.M., and K. Chandra. 2007. Energy-based equivalence between damage and fracture in concrete under fatigue. *Engineering Fracture Mechanics* 74: 2320–2333.

8. Song, Y.P. 2002. *Constitutive relation and failure criterion of a variety of concrete materials*. Beijing: China Water Conservancy and Hydropower Press.

9. Gokce, A., S. Nagataki, T. Sseki, et al. 2004. Freezing and thawing resistance of air-entrained concrete in-corporating recycled coarse aggregate: The role of air content in demolished concrete. *Cement and Concrete Research* 34 (5): 799–806.

10. Zhao, J.K. 2009. 3D meso-scale failure simulation of four-phase composite concrete. *Journal of Civil, Architectural & Environmental Engineering* 31 (4): 37–43.

11. Zhao, J.K., and J.J. Wen. 2013. Progressive failure process of rock excavation based on meso-mechanical method. *Journal of Disaster Prevention and Mitigation Engineering* 33 (3): 316–322.

12. Zhao, J.K., S.H.J. Huang, D. Wang, et al. 2015. Research on Rainfall-induced Debris Flows and Soil Particle Movement. *Chinese Journal of Soil Science* 46 (5): 1089–1095.

13. Sain, Trisha, and J.M. Chandra Kishen. 2007. Energy-based equivalence between damage and fracture in concrete under fatigue. *Engineering Fracture Mechanics* 74: 2320–2333.

Influence of Working Vehicles on Traffic Operation in Regional Road Networks Based on Microscopic Traffic Simulation

Jiaqi Shang, Xuedong Yan and Jinxian Weng

Abstract With the development of urban transportation, more and more working vehicles appear on the road network for maintenance, cleaning, dust suppression etc. However, working vehicle would produce the negative effects to the regional road network when providing a safety and comfortable environment to drivers. Therefore, this paper would conduct the research about the influence of working vehicle on region road network by using microscopic simulation. An influence analysis model based on microscopic simulation will be introduced which contains Paramics Modeler Module and Data Analysis Module. This study focus on the research that how the different working strategy of working vehicle (travel speed, travel lane and working start-time) give an effect on the whole delay of vehicles on the region road network and road safety under the given operation path through importing the control plug-ins of working strategy, which developed with the Application Program Interface (API) in Paramics. The research shows that the influence caused by working vehicle is related to the traffic volume on region road network during the working time. Moreover, the influence on the aspects of the whole delay and links' safety is different when working vehicle implements the task with different working strategy. Higher speed on the first lane or lower speed on the second lane would produce a big influence on region road network, which is especially obvious in early peak or latter peak. The paper would illustrate the mechanism of the action between working strategy and the influence caused by the working vehicle. Besides, the result of this research will help to formulate a better working plan in reducing the negative effects of working vehicle on road network.

J. Shang · X. Yan (✉)
MOE Key Laboratory for Urban Transportation Complex Systems Theory
and Technology, Beijing Jiaotong University, Beijing 100044, China
e-mail: xdyan@bjtu.edu.cn

J. Shang
e-mail: 14120874@bjtu.edu.cn

J. Weng
College of Transport and Communications, Shanghai Maritime University,
Shanghai 201306, China
e-mail: wjx206@gmail.com

Introduction

With the development of urban transportation, road traffic environment has gained more and more attentions. In order to ensure that the drivers in the process of driving can enjoy a clean, comfortable and safe traffic environment, road operations implemented by working vehicle is needed frequently. The most common working vehicles, such as road sweeper, sprinkler, and snow clearer, are playing an important role in urban network. The task of urban road maintenance carried out by the working vehicle contributes to beautifying the road environment and the improvement of traveling environment. However, it is inevitable that the working vehicle can also bring some negative problems when carrying out the working task. The task would produce the influence on the whole delay of vehicles on road network, accident risk of links on the working path and interference to running state of the road network. These problems will give an impact on the service level of road network. Thus, it has become a focus that how working vehicle carries out the working task in order to reduce the interference to running state of the road network so that the road network keeps relatively high level of service and security.

As we all know, working vehicle cannot travel randomly on the road just like other common vehicles because its task decides it has its own driving characteristics. In order to finish the task with high quality, the working vehicle needs to keep driving at a constant speed as much as possible and cannot change lanes discretionarily. Due to the characteristics of working vehicle, the research of the influence caused by working vehicle with diffident working strategy is significant which contains working start-time and travel characteristics of speed and lane. Meanwhile, the influence of working vehicle on the region road network is different from its on network.

Therefore, the research will be proceeded from the perspective of region road network. The paper not only considered the total delay of the whole region road network, but introduced the variance of links on path as an index to evaluate the safety of links. Besides, the relationships between the travel characteristic of lane and speed, working start-time and the influence will be studied by comparing different working strategy. This paper will introduce an influence analysis model based on microscopic simulation and then a comprehensive understanding will be acquired about the influence of working vehicle from different perspective, which will help reduce the disadvantage of the influence and improve road safety.

Literature Review

Currently, road working includes two forms: fixed working and mobile working. Now, the researches of road working are focused on fixed working zone. In the research on the influence of construction strategies for work zone, Zheng et al. [1] used the method of evaluating network-wide traffic delay to analyze drivers' behavior for the several construction strategies in the research of measuring network-wide traffic delay in schedule optimization for work zone planning in urban networks. And on this basis, simplified methods were developed by Qu [2] which identify and model the work zone impact on travel time and travel time reliability to provide evaluation input for the work zone scheduling problem. In the research field of the influence caused by woke zone in the period of working, the research of work zone traffic analysis and impact assessment was studied by Schroeder et al. [3] which focus on the development and calibration of a network-wide microscopic simulation model to predict the network-wide impacts of this work zone during construction. In order to reduce the impact on the network from work zone, many scholars had done the researches about what strategies need to be taken to improve the running state of vehicles in working zone. Kim et al. [4] developed a simple and practical work zone simulation model to provide a decision-making framework in the research of a decision-making framework to evaluate traffic control strategies on freeway work zone, which can be used to evaluate various work zone traffic control strategies. Besides, a methodology was developed by Morgado et al. [5] also described to support decision makers in planning work zones in the research of work zone planning in pavement rehabilitation: integrating cost, duration and user effects.

Meanwhile, there are also some scholars of giving researches about mobile working vehicle, which focus on the working effect of working vehicle. Many researchers (e.g., Amato and Querol [6] and Chang et al. [7]) conducted the research about the effectiveness of preventive measures on improving air quality such as street sweeping, washing and dust suppressants. In addition to the research of road sweeper and sprinkler, there are also some scholars carrying on the research of winter snow shoveling. Yu et al. [8] did the study of optimization of snow emergency vehicle scheduling under real-time information and established an integer programming model of the snow emergency vehicle to increase the level of winter road maintenance and lay a foundation for practical applications. In order to be able to formulate the optimal shoveling schedule, the breadth-first search algorithm was adopted by Wang et al. [9] to optimize the scheduling of snow removers in the research of optimal scheduling and simulation of snow removers, which showed that the efficiency of cleaning work can be improved and the delayed time of the airport open can be decreased. As for the influence of such working vehicle on transport systems' operation, Zmuda-Trzebiatowski [10] made the study of selected aspects of road cleansing in the city of Poznan, which presenting the road cleansing goals, its influence on transport systems' operations and analyzing decision problems of similar areas such as road maintenance and transportation.

Besides these researches, some scholars were devoted to studying a method which could represent the safety of roads. Qu [11] showed that speed dispersion was a better predictor of the two microscopic potential risks and is recommended to proactively assess road safety because it provides consistent risk evaluation with microscopic potential risks. In the research of the relationship between speed variance and safety, speed variance and its influence on accidents were studied by Garber and Gadirau [12], which found that accident rates do not necessarily increase with increase in average speed but do increase with increase in speed variance. In its research, it showed that the speed variance could better indicate the road safety. As for the choice of evaluation index in research of network, Jenelius and Koutsopoulos [13] presented a statistical model for urban road network travel time estimation and in its paper traffic delay was used as an evaluating indicator to optimize in urban network in its research. In such studies, the simulation was also used to acquire the travel time and make quantitative analysis. A multi-criteria decision support methodology was used by Yang and Regan [14] for implementing truck operation strategies through the microcosmic traffic simulation software and acquired various performance measurements that can describe quantitative impacts. It chose the travel time as a variable to give an analysis on urban road networks.

But the researches rarely mentioned the influence on vehicles in delay and safety caused by the working of sweeping, washing and dust suppressants and researchers did not provide a good method to conduct the research about the influence on network caused by working vehicle. Therefore, this paper would introduce a better performances and comprehensive method to the research based on the previous research.

Objectives and Contributions

The objective of this paper is to develop an influence analysis model based on microscopic simulation to accurately estimate the influence caused by working vehicle when it carries out the working task. The influence of working strategy varies with working start-time, travel speed and travel lane. Therefore, this study contributes to implement the research by using method of simulation.

To achieve this objective, the variables of working start-time, travel speed and travel lane are used to restraint the working strategy of working vehicle. The variables of the whole delay of vehicles on the region road network are used to quantify the influence caused by working vehicle. The study makes the best use of API interface function in Paramics. The control plug-in was created to achieve the control of working strategy including working start-time, travel speed and travel lane. Paramics simulation was applied to estimate the influence on the region road network which would vary with the different working strategy. The working of dealing with the data of different working strategy should be done to interpret the mechanism of interaction between the working vehicle and the influence.

A relatively complete influence analysis model based on microscopic simulation was built to provide a research method for the study of similar problems.

Research Method

Research Approach

The influence of working vehicle on the regional road network mainly contains two aspects. First, working vehicle will have an effect on the delay of vehicles in region road network. Secondly, working vehicle also will give an influence on the safety of links composing the path of working task. Therefore, the model will contain two parts that are analysis of the whole delay collected from all vehicles in region road network and analysis of safety about links. As for the safety of links, the maximum of mean velocity variance of every 5 min in influential time range will be used as link safety indicators to be analyzed, which can reflect link safety. If it is not especially pointed out in this paper, link speed variance maximum is referred to the maximum of mean velocity variance of every 5 min in influential time range.

The research of this paper will be conducted as below:

Step 1: Obtain the data of traffic volume, links attributes, intersection signal timing etc. about the network of research;

Step 2: Reasonable layout OD according to the information of network;

Step 3: Build the road network model in Paramics;

Step 4: Develop the plug-ins of the control of working vehicle's behavior by using API in Paramics;

Step 5: Simulate the model and Data output;

Step 6: Data Analysis and Processing;

Then, in order to prove the availability of the method, a practical region road network of Haidian District in Beijing was selected as the object of study in order to objectively reflect the influence of working vehicle. The research will be conducted by comparing with simulation data from different working strategy. We will illustrate its internal influence mechanism from the microscopic view by analyzing the relationship between the influence level and the three variables that includes working start-time, speed and lanes. The research will give us deeper insight about the delay of vehicles and the safety of links.

Simulation Methodology

In the real case, the whole delay of vehicles on region road network and the speed of vehicles on links are hard to acquire, which brings great challenge to the research. But the emergence of the microscopic simulation software supplies us

with a kind of new perspective. The research of this paper is implemented by means of using Paramics, a microscopic simulation software. The software itself has a powerful API interface and huge function library which enable users to develop the control plug-ins with the API interface according to the need for the research. The plug-ins are able to control the working vehicle according to the research objective. Therefore, this paper will introduce an influence analysis model based on microscopic simulation to conduct the research (Fig. 1). The model includes two parts: Paramics Modeler Module and Data Analysis Module. Paramics Modeler Module contains the modeling of regional road network and the control of behaviors of working vehicle. Data Analysis Module contains the compute of the whole delay of vehicles on region road network and the maximum of 5-min-average road speed variance on the working path.

Paramics Modeler Module

The study of this paper will need to use three modules of microscopic simulation software Paramics. The three modules are modeller module, programming module and processor module. Modeller module is mainly used to implement the modeling of the road network. Programming module is mainly used to realize the control of working strategy. The processor module is used to realize the batch of simulation and high speed simulation to have faster access to the simulation data.
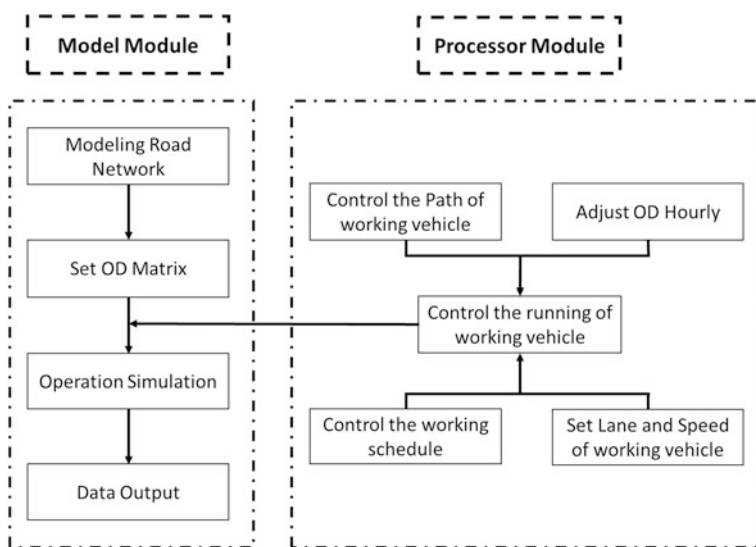
Fig. 1 Conceptual procedure for simulation model for impact analysis

Data Analysis Module

The whole delay was acquired by comparing the two situations whether there was a working vehicle during the period of research. The difference is the whole delay of vehicles. Because of the influence caused by working vehicle, the count of vehicles collected by simulation on road network would not be equal within the same period of the study time. But the difference was very small between the two situations. So the data of whole delay would be amended slightly in order to make the data of delay more accurate. In the hypothesis of having a working vehicle, the whole existence time of vehicles collected by simulation on road network is t_Y and the count of vehicles is n_Y. In the hypothesis of having no working vehicle, the whole existence time of vehicles collected by simulation on road network is t_N and the count of vehicles is n_N. The corrected formula of the whole delay is below:

$$\text{Delay} = (t_Y - t_N) + (n_N - n_Y) \times t_N / n_N \tag{1}$$

As known above, link speed variance is the maximum of 5-min-average road speed variances on the working path within the period of the study time. Assumed that the count of vehicles on road is n_{RV}, the speed of vehicles returned from the simulation in every second is v_i $(i = 1, \ldots, n_{RV})$. The formula of link speed variance is below:

Average link speed:

$$v_{as} = \sum_{i=1}^{n_{RV}} v_i / n_{RV} \tag{2}$$

Five minutes average road speed variance:

$$\text{var}_n = \sum_{t=1}^{300} \sum_{i=1}^{n_{RV}} (v_i - v_{as})^2 / 300 \tag{3}$$

The maximum of 5 min average road speed variance:

$$\text{variance} = \text{Max}(\text{var}_1, \text{var}_2, \text{var}_3, \ldots, \text{var}_n) \tag{4}$$

Through the combination of modeling module and the data analysis module, the model not only considered the influence of delay from the perspective of travel costs, but also included the analysis of the accident risk caused by working vehicle from the perspective of safety. Therefore, the model provided the researchers with a good method to do the similar research.

Case Study

Network Descriptions

BJHDRRN is a practical region road network of Haidian District in Beijing. Figure 2 is the general situation of the road network. Based on investigation of traffic volume on the region road network, traffic flow data from 12:00 to 13:00 were acquired. Table 1 is the basis matrix of traffic volume. Then, we got the traffic flow data from 6:00 to 22:00 that were calculated with the coefficient of traffic hours by consulting the related literature. Figure 3 is the hour factor of traffic volume variation trend.

Considering the road condition, the travel lane has two choices. These two choices are the first lane and the second lane. The first lane refers to the lane that is close to the edge of the road. The second lane refers to the lane that is next to the first lane. In this paper, if it is not particularly pointed out, lane = 1 refers to the lane that is close to the edge of the road and lane = 2 refers to the lane that is next to the first lane. According to the working task, the travel speed of working vehicle can
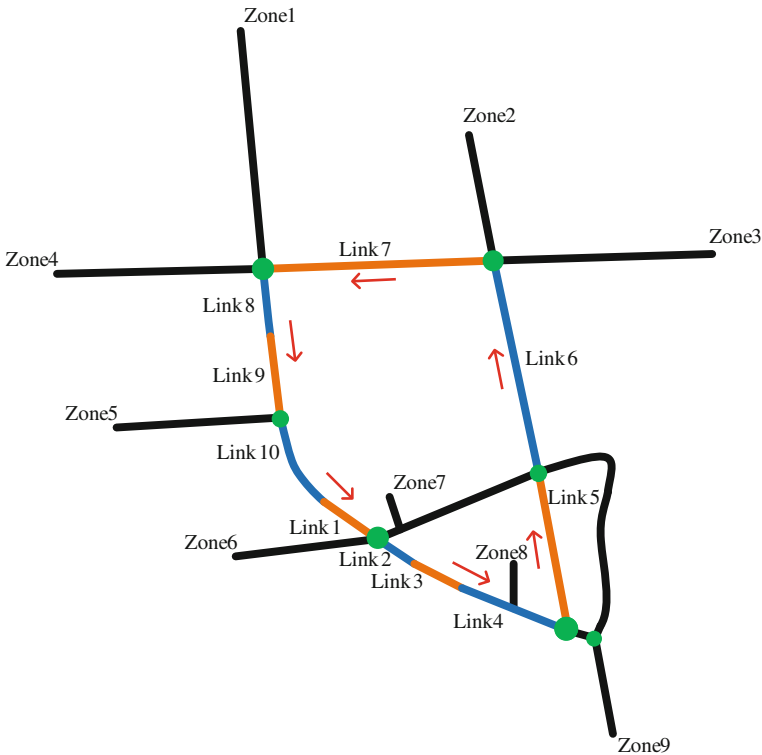
Fig. 2 Overview of the region road network

Table 1 Basis OD matrix of traffic volume

O	D									
	Zone 1	Zone 2	Zone 3	Zone 4	Zone 5	Zone 6	Zone 7	Zone 8	Zone 9	Total
Zone 1	0	80	150	100	125	90	20	10	150	725
Zone 2	50	0	40	65	55	75	25	20	50	380
Zone 3	100	80	0	160	60	90	25	20	130	665
Zone 4	100	40	150	0	90	140	25	25	130	700
Zone 5	40	25	25	30	0	45	20	20	60	265
Zone 6	80	40	50	75	40	0	15	20	90	410
Zone 7	25	25	20	25	25	25	0	10	35	190
Zone 8	20	25	20	20	15	20	10	0	40	170
Zone 9	120	80	100	120	100	80	20	15	0	635
Total	535	395	555	595	510	565	160	140	685	4142
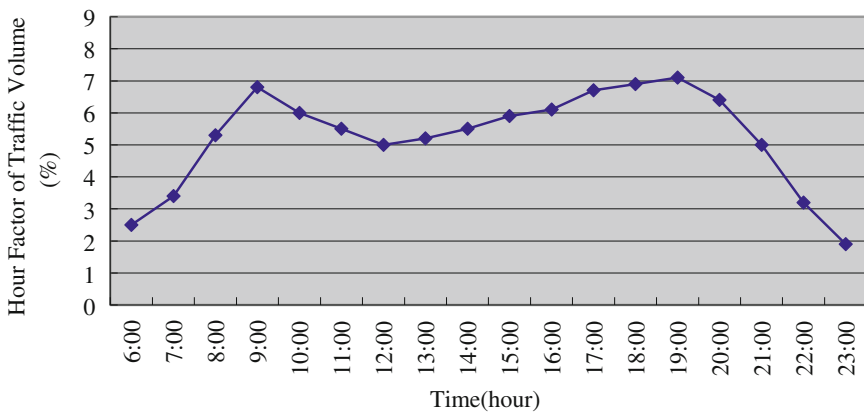
Fig. 3 Hour factor of traffic volume variation trend

adopt two choices 10 and 20 km/h to facilitate comparative analysis. In order to study the influence of the operation scheme at different times of the day, 32 operating schemes at the different start-time throughout the daytime are selected to analyze the influence of the operation scheme.

As for safety analysis, six links were selected in the working path to show the influence on links' safety caused by working vehicle. The working path consists of the colored lines in Fig. 2. The reason why we chose the six links was that this six links was representative which including three links of two lanes, one link of three lanes and two links of four lanes. Table 2 is the attribute of roads on the working path. Therefore, the result would show the inherent relationship between the working start-time, travel speed and lane of working vehicle and the influence caused by the working task.

Table 2 Attribute of links on the working path

Link	1	2	3	4	5	6	7	8	9	10
Number of lanes	4	2	3	4	2	2	2	3	4	4
Basis traffic volume (veh/h)	685	615	615	505	180	360	490	920	920	685

Network Traffic Delay Analysis

In the research of influence on the regional road network, the total delay of whole vehicles on the region road network and link speed variance maximum are chosen to express the level of influence. So we can give a quantitative evaluation about the influence of working vehicle on the region road network. The range of working time in this paper is from 6 am to 10 pm. The time point of starting working task is at 15th minute and 45th minute past every hours. The working vehicle with different working strategy was simulated to acquire the data about delay and road speed variance.

As for every working strategy of working vehicle when we collected the data of delay including whole vehicles on the region road network, we considered the influential time range of working vehicle as 2 h latter from starting the task instead of the period working vehicle just spent on. We chose 2 h as the period of collecting data because the data of delay maybe present the deviation when the influential time range of collecting data was longer or shorter. If we choose shorter period of influential time range about the working vehicle, the influence of working vehicle would be not acquired completely. Likewise, the delay caused by the influence of working vehicle may be concealed by other factors and may even be overwhelmed if the period was longer, which would make the research become meaningless. So we set the period of collecting delay data as 2 h.

We already mentioned above that the three variables are the main focus of our research containing working start-time, travel speed and lane of working vehicle. The motion of working vehicle during operation is usually at 5–25 km/h. So we choose 10 and 20 km/h as the speed we studied. The first lane and the second lane were selected as the variable factors. A large number of data of whole delay and road speed variance maximum would be got by simulating different operation solutions.

Traffic Delay Caused by Working Vehicle's Speed and Driving Lane

There are four travel characteristics we carried out in this paper. They are lane = 1, speed = 10 km/h; lane = 1, speed = 20 km/h; lane = 2, speed = 10 km/h; lane = 2, speed = 20 km/h.

The research indicated that the influence level caused by the working vehicle with different working strategy was different. The influence level is related with the traffic volume on the region road network. Besides, when working vehicle travels
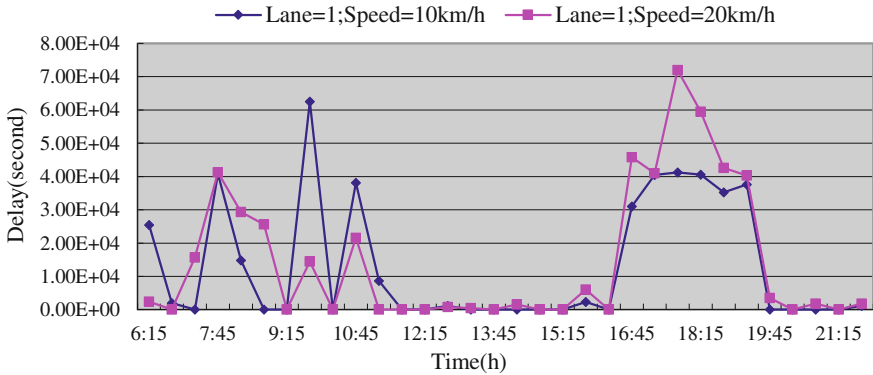
Fig. 4 Delay of different speed on the first lane at different working start-time

on different lane and adopts different speed, the influence level would be discrepant even if the working vehicle conducted the working task at the same time.

The research shows that the influence level of delay posed by the speed 20 km/h of working vehicle is relatively higher in latter peak when the task was implemented on the first way, as it is shown in Fig. 4. But when the start-time of working was in the morning, this tendency is not obvious. There just are few working start-time at which the working vehicle with speed 10 km/h produced higher influence of delay. This may be due to the speed of change on traffic volume. On the whole, the speed of 20 km/h would bring higher influence of delay than 10 km/h when working vehicle travels on the first lane. This phenomenon is mostly because higher speed on the first lane may result in greater interference to the running state of road network compared with lower speed.

However, the influence level of delay caused by different speed was different when the working vehicle traveled on the second lane. The lower travel speed produced bigger influence on the region road network. Owing to the speed of vehicles traveling on the second lane is higher, the lower speed of working vehicle would produce huge interference to the traffic flow of this lane compared with higher speed. This is due to the fact that some vehicles on the second lane need to slow down to merge into the first lane for completing the overtaking maneuver when vehicles would like to overtake if the road section only has two lanes. If the behavior happened, it would seriously increase the delay of vehicles on the second lane. Meanwhile, vehicles on the first lane may suffer more interference. As it was showed in the Fig. 5. When the working vehicle conducts the task on the second lane, the speed of working vehicle is sensitive to the whole delay of vehicles on region road network.

Likewise, the influence level of delay produced by the task needed to complement on the second lane was higher compared with the task needed to finish on the first lane when working vehicle traveled at the lower speed especially in late peak. Just as Fig. 6. The vehicles traveling fast relatively on the second lane may suffer
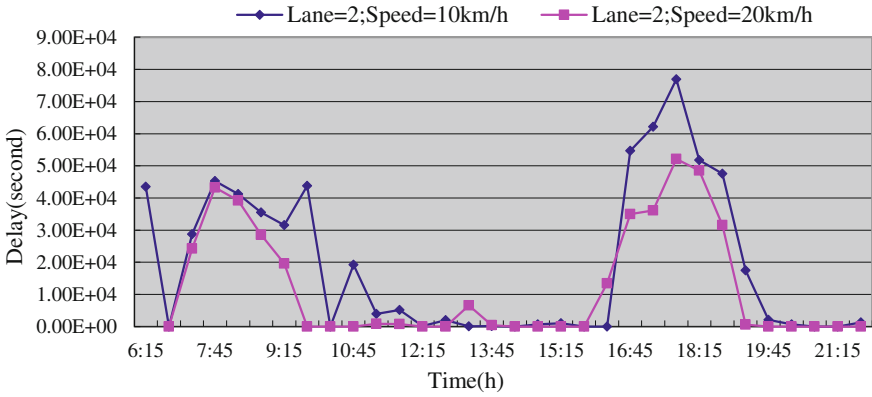
Fig. 5 Delay of different speed on the second lane at different working start-time
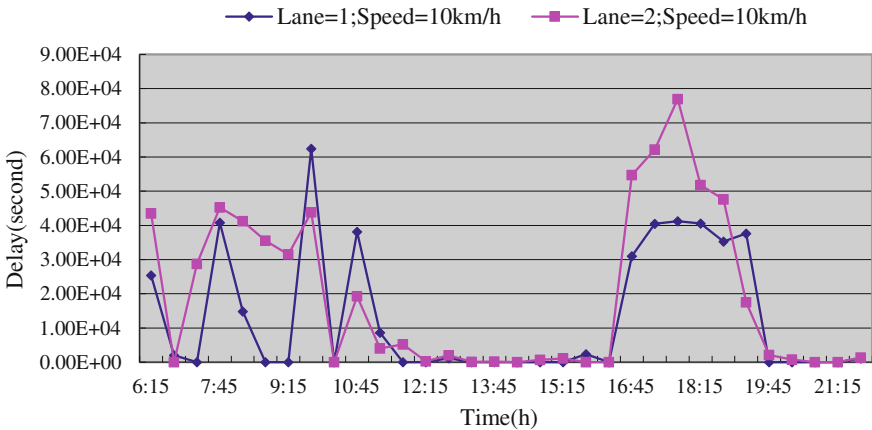
Fig. 6 Delay of different lane with the speed of 10 km/h at different working start-time

more interference when working vehicle travels with lower speed on the second lane, which would bring huge delay of vehicles on the second lane. But when working vehicle travels carries out the task on the first lane, the second lane suffers less interference due to the lower speed of working vehicle and the vehicles traveling on the first lane need to increase the speed in order to adapt the relatively fast traffic flow on the second lane to complete the overtaking maneuver. Therefore, the result is that the task implemented on the second lane result in higher delay than the first lane when working vehicle in lower speed.

Conversely, the influence level of delay would be relatively higher in the first lane when the working vehicle traveled at higher speed. Although the interference produced by working vehicle would decrease with the increase in working speed,
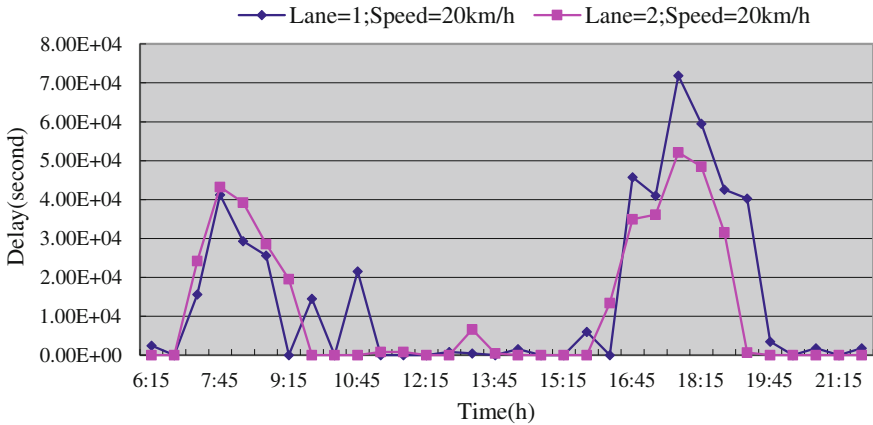
Fig. 7 Delay of different lane with the speed of 20 km/h at different working start-time

the second lane of relatively faster speed is more sensitive to change of working speed. So the change of interference decrease caused by the working task on the second lane is larger compared with the task on the first lane when the working vehicle traveled at higher speed. Therefore, there is nothing to strange that the delay caused by the working task on the first lane is larger as Fig. 7 shows.

In other words, we should avoid to operate at higher speed when working vehicle implements the task on the first lane. Similarly, we should avoid to operate at lower speed when working vehicle implements the task on the second lane. The difference of lane that working vehicle travels on will bring higher delay than the difference of speed to the region road network.

Traffic Delay Caused by Working Vehicle's Operation Time

In order to find out the relationship between the whole delay of vehicles and working start-time, we stacked the delay of four travel characteristic by dealing with the simulation data. We found that the relationship between the whole delay of vehicles and traffic volume on the region road network were very close no matter what characteristics the working vehicle with for a certain task. It also related to the change speed of traffic volume. As the Fig. 8 shows, the working task would produce the serious influence on the whole delay of vehicles when the working task is carried out under the condition of larger traffic volume, such as earlier peak or later peak in a day.

Research shows that the task carried out in later peak would produce higher influence than the task carried out in earlier peak. Besides, the range of time that would cause the influence is concentrated in later peak. But the relevance between the influence and the working time is not obvious in the morning. The result may be caused by the change speed of traffic volume, and there is a need to give a further
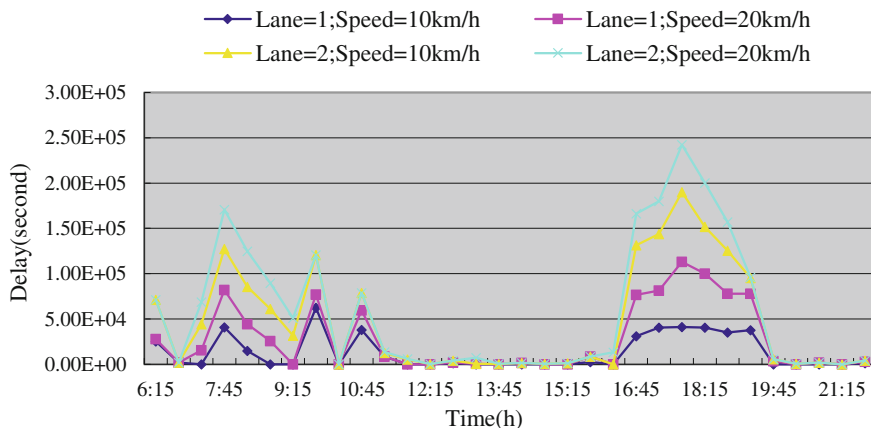
Fig. 8 Accumulation of the delay of four different travel characteristics

study about the specific relationships between the change speed of traffic volume and its influence. We also found that the working vehicle almost made no influence on the region road network in off-peak hours, which traffic volume is lower.

Road Segment Safety Analysis Based on Speed Variance as Surrogate

The influence of working vehicle is not only embodied in the delay of vehicles, but also in the road safety of working path. As for the safety influence produced by the working vehicle, we used a surrogate method in the influence analysis model based on microscopic simulation. It is link speed variance. The road safety would be evaluated by analyzing the road speed variance data.

In this paper, the road speed variance maximum refers to the maximum of 5 min mean in 1 h. Here, we did not choose 2 h as the period of collecting data just as we did in the analysis of delay because this method would make the variance data better performance on reflecting the road safety when working vehicle pass. First, we decided to set 1 h as the period because the duration time of working is relatively short. If we set 2 h as the period of collecting data, the road speed variance data we need is likely to be covered with the biased road speed variance data Thus, the link speed variance data cannot well reflect the road safety. Besides, due to the change of daily traffic, the research range of 1 h could avoid the effect on the decrease in road safety result from the increase or change in traffic volume. Therefore, we choose 1 h as the time range of research. So next the six representative links on the working path would be used to show the influence of different working strategy on link safety. The number of lanes and basis traffic volume about the six road segments are shown in Table 3.

Table 3 Attribute of six links on the working path

Link	4	5	6	7	8	10
Number of lanes	4	2	2	2	3	4
Basis traffic volume (veh/h)	505	180	360	490	920	685
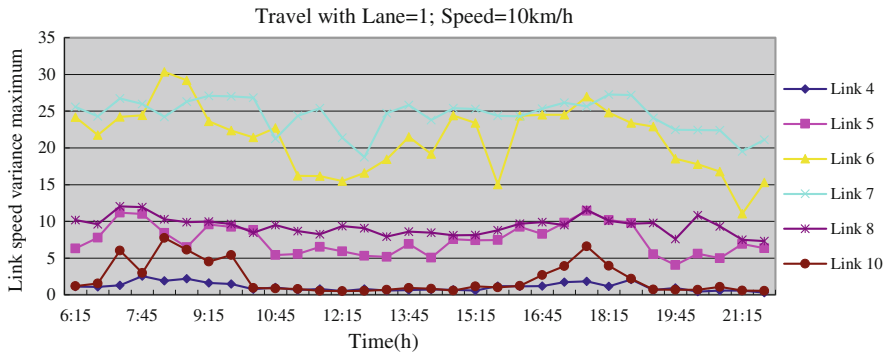
Fig. 9 Roads speed variance maximum under the through of working vehicle with lane = 1; speed = 10 km/h

Analysis of Six Representative Links Safety Under Different Working Strategy

The research found that road safety was different when the working time was different. The variation tendency of road accident risk is consistent with the trend of daily traffic. Road safety was also slightly different when the working vehicle carried out the task with different travel characteristic even if the working time was same.

The data shows that the change of daily traffic has bigger influence on the road safety. The accident risk of road on the working path is different even if the travel characteristic of working vehicle is same when working start-time is different. The change of link 4, 5, 6 are most obvious. Meanwhile, the accident risk would increase with the increase in traffic volume as the link 5, 6, 7 show under the different traffic volumes and same lanes showed in Figs. 9, 10, 11 and 12. As for a working strategy, namely the working start-time and travel characteristic are known. The accident risk of different roads on the working path is also different, which is formed by the different attribute of road. The difference of roads accident risk is related to the factors such as the number of lanes on link and traffic volumes on the link. Under the low and medium traffic volume, higher traffic volumes or fewer lanes would bring about higher accident risk. The roads accident risk are the most highest in early peak or latter peak. Roads accident risk caused by working vehicle is not same when the travel characteristic of working strategy is different. So it is essential to choose appropriate working time in a day to carry out the different working task. Besides, the working task should avoid be scheduled in peak so as to not bring higher accident risk to the region road network.
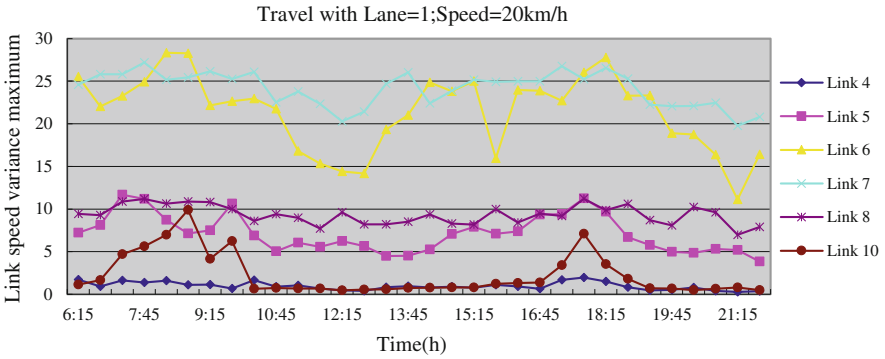
Fig. 10 Roads speed variance maximum under the through of working vehicle with lane = 1; speed = 20 km/h
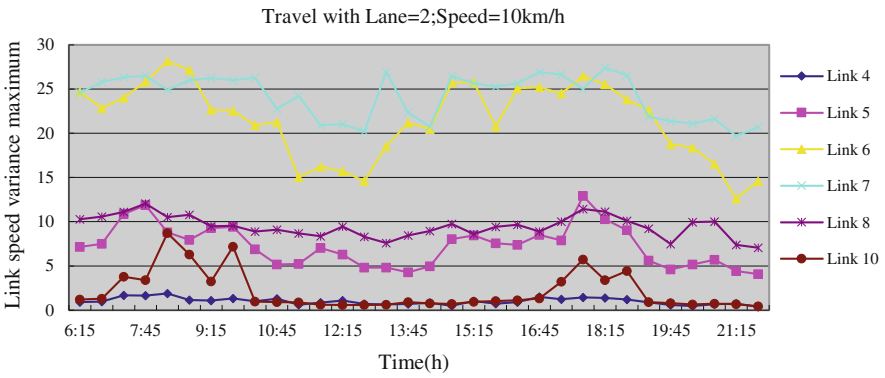
Fig. 11 Roads speed variance maximum under the through of working vehicle with lane = 2; speed = 10 km/h
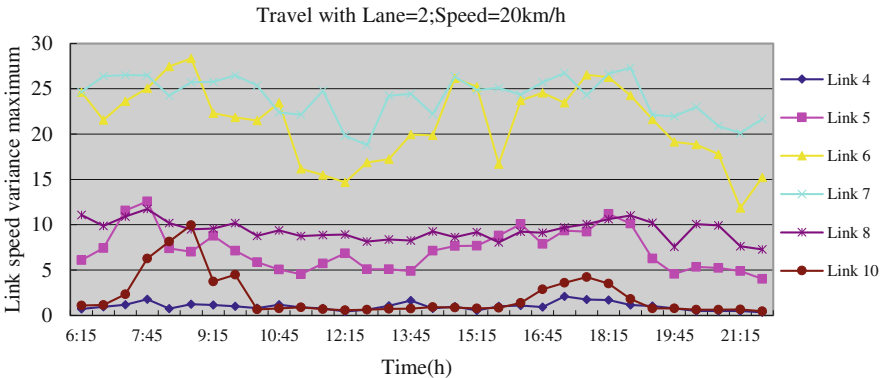
Fig. 12 Roads speed variance maximum under the through of working vehicle with lane = 2; speed = 20 km/h

The Difference Analysis of Links Accident Risk Compared with no Working Vehicle

In the research of roads accident risk caused by working vehicle, three representative links with different number of lanes were used to analyze the relationship between the accident risk and travel characteristics of speed and lane, working start-time. Link 6 is a four lanes wide road. Link 8 is a six lanes wide road. Link 10 is a eight-lane-wide road. The data show in the form of figures as below by contrasting the accident risk of links on the working path under the state that whether there is a working vehicle. As for working start-time, the relationship between the accident risk and working start-time is the relationship between the accident risk and traffic volumes in nature. When working vehicle travels with a certain characteristic, the accident risk of road on the working path caused by the working task is different when working start-time is different. The reason is that traffic volumes change. Meanwhile, the roads that working vehicle does cause the influence of accident risk on are changed, when travels characteristic changes. On the whole, the accident risk caused by the working task is higher when lanes of link were fewer, compared with the working task implemented on more lanes of link. The research also found that a working task was not producing the accident risk to all roads and maybe it just caused the accident risk to some roads on the working path as it showed in Figs. 13, 14, 15 and 16.

The research found that the influence level on link safety caused by the working vehicle was different when working vehicle traveled with different speed or lane. As showed in the four figures, the influence on link safety caused by working vehicle is different among the three different number of lanes when traffic volume are not large in network. Link 6 suffered the highest accident risk. Link 10 suffered the smallest accident risk which was just higher in early or latter peak and almost would not be affected in off-peak traffic. The reason of this result is that the number of
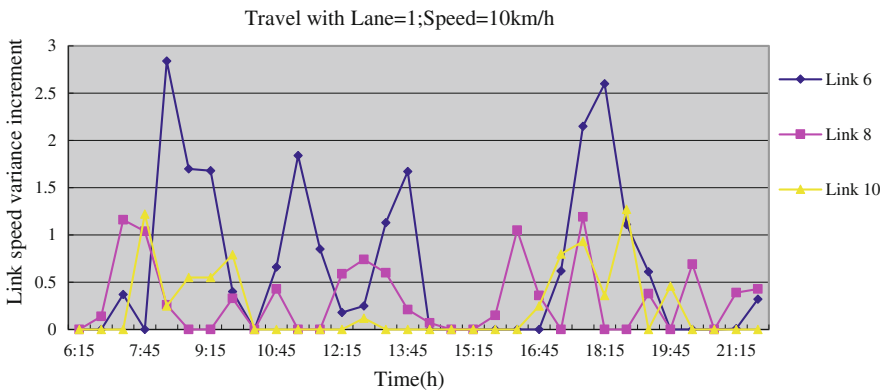
Fig. 13 Difference of links accident risk caused by working vehicle with lane = 1; speed = 10 km/h
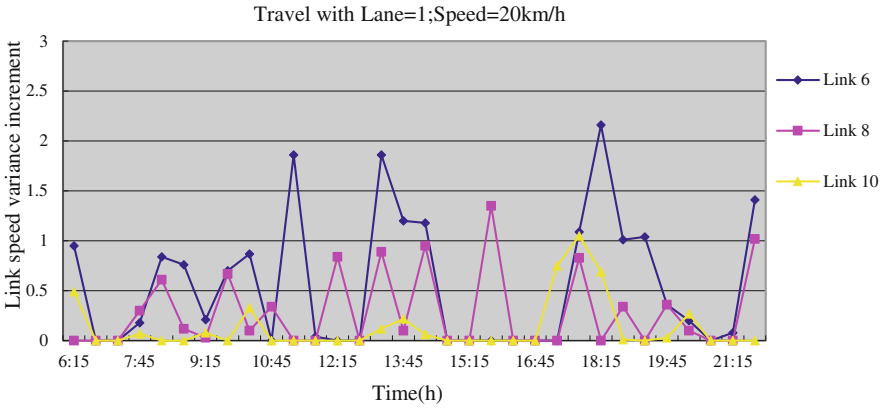
Fig. 14 Difference of links accident risk caused by working vehicle with lane = 1; speed = 20 km/h
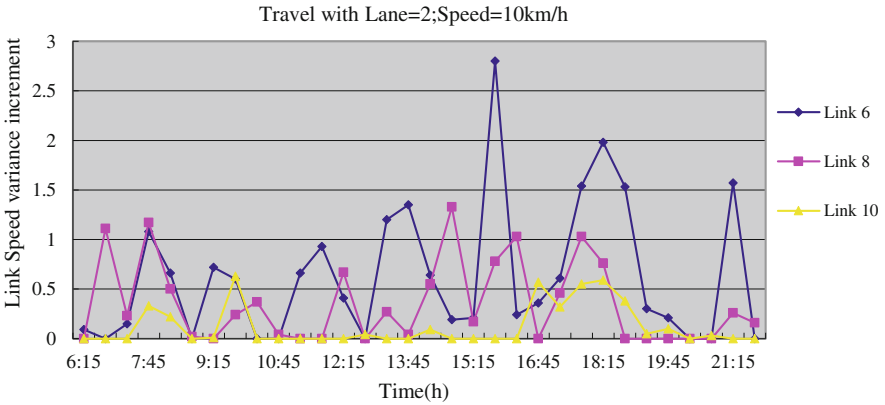
Fig. 15 Difference of links accident risk caused by working vehicle with lane = 2; speed = 10 km/h

lanes is different. The higher number of lanes in a link, the less influence on link accident risk. The change of the influence on link 8' accident risk was approximately serrated with working start-time changed, which was affected by the different traffic volume at 15th minute and 45th minute past every hours. Besides, the accident risk of roads would increase when the working task was implemented in early peak or latter peak.

However, the influence on accident risk for every link varied with different characteristics of lane and speed, which is related to the attribute of link such as the number of lane and traffic volumes on the link. The working vehicle with different travel characteristics has different influence on every road section even if the working time is same as figures show above. When working vehicle travels with 10
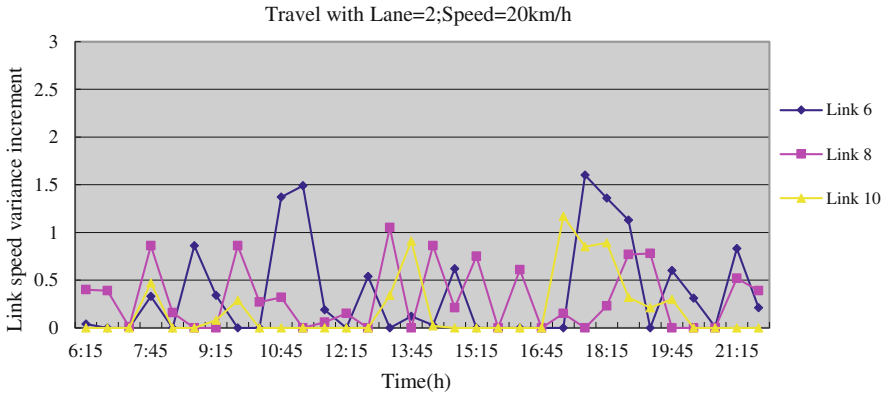
Fig. 16 Difference of links accident risk caused by working vehicle with lane = 2; speed = 20 km/h

and 20 km/h, the data showed that the accident risk of 10 km/h is higher because the speed of 10 km/h working vehicle traveled with is lower compared with the average speed of link. So the speed of 10 km/h would produce higher accident risk. When the working task was conducted on the second lane, the change of speed would produce bigger difference on the accident risk of roads.

But it is worth mentioning. The research also shows that the working task at a certain time would not cause influence on all links. We learned that the existence of working vehicle was not always producing the bad effects to the links on the region road network. In some cases, the existence improves the road safety on the contrary. This is not common in the last study. In all, by comparing the data of links speed variance whether there is a working vehicle, it shows that the influence of road safety caused by working vehicle is very complicated. The road safety was affected by many factors, such as the change of daily traffic, working start-time, travel characteristic and attribute of links.

Conclusion

This paper introduced an influence analysis model based on microscopic simulation. Besides, the road network model was built and the control plug-ins of working strategy were developed. Based on the model, the research was conducted by analyzing the influence of the whole delay and road safety on network result from different working strategy of working vehicle by simulation.

Through the research of this paper, we can acquire about the influence on region road network produced by the working vehicle. The influence of delay on region road network caused by working vehicle has close links to working start-time and travel characteristic. There would produce a big influence (i.e., delay and safety) on

road network when the working vehicle travels on the first lane at a higher speed or lower speed on the second lane. This phenomenon is relatively obvious in peak hour. As for the influence of working start-time, it shows that the change of traffic volume brings impact to vehicles on the regional road network in nature. Although the road safety was affected by the change of daily traffic, working start-time, travel characteristic and attribute of links, the working vehicle is not always bringing bad effects to the road safety.

The analysis result on the specific research object may be a little different when research object changes. The change of research object may be reflected in region road network scale, attribute of lane and the variation of daily traffic volume changes. But the method of influence analysis model based on microscopic simulation possesses favorable commonality. The influence analysis model based on microscopic simulation can better reflect the influence caused by the working vehicle, which provides a good method to deal with the similar issues. In addition, the methodology used in this study would provide an effective way to identify the best alternative among potential working strategies for a wide range of transportation systems management applications. However, this paper only did the research about working start-time and travel characteristic. In fact, the influence on region road network caused by the working vehicle is also affected by other factors such as attribute of links, network scale and the speed of change on traffic volume. These factors need to further study in the future.

References

1. Zheng, Hong, Eric Nava, and Yi-Chang Chiu. 2014. Measuring network wide traffic delay in schedule optimization for work-zone planning in urban networks. *IEEE Transactions on Intelligent Transportation Systems* 15, no. 6, 2595–2604.
2. Qu, Tao, et al. 2015. Travel time reliability based highway work zone scheduling. Transportation Research Board 94th Annual Meeting. No. 15-3933.
3. Schroeder, Bastian J., et al. 2014. Work Zone Traffic Analysis & Impact Assessment. No. FHWA/NC/2012-36.
4. Kim, Taehyung, Taehyeong Kim, and Hyoungsoo Kim. 2013. A decision-making framework to evaluate traffic control strategies on freeway work zone. *International Journal of Control and Automation* 6, no. 2: 45–52.
5. Morgado, João, and Jose Neves. 2014. Work zone planning in pavement rehabilitation: Integrating cost, duration, and user effects. *Journal of Construction Engineering and Management* 140, no. 11: 04014050.
6. Amato, F., and X. Querol. 2010. A review on the effectiveness of street sweeping, washing and dust suppressants as urban PM control methods. *Science of the Total Environment* 408: 3070–3084.
7. Chang, Yu-Min, et al. 2005. Effectiveness of street sweeping and washing for controlling ambient TSP. *Atmospheric Environment* 39, no. 10: 1891–1902.

 8. Yu, Yang, Gang Liu, and Jian Wang. 2009. Optimization of snow emergency vehicle scheduling under real-time information. *Science Technology and Engineering* 19: 024.
 9. Wang, Li Wen, et al. 2011. Optimal scheduling and simulation of snow removers for airfield pavement based on breadth-first search. *Advanced Materials Research* 214.
10. Zmuda-Trzebiatowski, Paweł. 2010. Selected aspects of road cleansing in the city of Poznan. *LogForum* 6, no. 3, 8.
11. Qu, Xiaobo, Yan Kuang, Erwin Oh. 2014. Safety evaluation for expressways: A comparative study for maroscopic and microscopic indicators. *Traffic Injury Prevention* 15: 89–93. (Taylor & Francis).
12. Garber, Nicholas J., and Ravi Gadirau. 1988. Speed variance and its influence on accidents. Reports—Research/Technical (143).
13. Jenelius, Erik, and Haris N. Koutsopoulos. 2013. Travel time estimation for urban road networks using low frequency probe vehicle data. *Transportation Research Part B* 53: 64–81.
14. Yang, Choong Heon, and Amelia C. Regan. 2013. A multi-criteria decision support methodology for implementing truck operation strategies. *Transportation* 40: 713–728.

Research on Comparison of Tram with BRT

Song Fang, Jianxiao Ma and Yancai Wang

Abstract Urban traffic is a combination of various kinds of transportation demand level. On the premise of the same kind of traffic demand, there may be many public transportation modes for policy makers to choose. To medium capacity of BRT and tram, this paper comparative analysis, respectively, from the aspects of capacity, speed, economical efficiency, and energy consumption. In conclusion, the average speed and capacity of both are almost, and the economical efficiency of BRT is little better than tram, while the energy consumption of tram is greater than BRT. On comprehensive consideration, it is suggested that policy makers give priority to tram under the same condition.

Keywords Capacity · Speed · Economical efficiency · Energy consumption

Introduction

With the development of economy, the traffic problem is becoming more and more prominent. In order to solve the traffic problems, many cities use the large volumes of public transportation. Big cities usually take subway or light-rail network planning. Due to the limited financial resources, the small cities generally selected tram or BRT. The tram and BRT systems are all belong to low and middle traffic volume of bus rapid transit system, and meanwhile, their function is pretty close. Comprehensive comparison of tram and BRT is shown in Table 1 . This paper will comparatively analyze them, respectively, from the aspects of capacity, speed, economical efficiency, and energy consumption.

S. Fang (✉) · J. Ma · Y. Wang
Automotive and Transportation Engineering Institute,
Nanjing Forestry University, Nanjing 210037, China
e-mail: 472916417@qq.com

S. Fang
Nanjing Institute of Railway Technology, Nanjing 210031, China

Table 1 Comprehensive comparison of tram and BRT

	Tram	BRT
Capability	6000–15,000 persons per hour	4000–12,000 persons per hour
Cost of vehicle	Importation: ¥25–30 million Domestic: ¥10–15 million	Domestic mainly: ¥1–2 million
Cost of construction	Ground form: ¥50–100 million/km Interchange form: ¥250–300 million/km	Ground form: ¥20–40 million/km Interchange form: ¥60–80 million/km
Energy consumption	0.07 kWh per sat non-pollution	0.28 kWh per sat CO_2 15 g per person per kilometer
Noise level db	70 db	75 db
Service period	20–30 years	10 years

Transportation Capability

Transportation capability refers to the maximum carried by vehicle in single direction peak hour. The calculating formula for the capacity of the trams is shown below:

$$C = P \times L$$

where

C is the largest capacity of tram in single direction, unit person/h;
P is a single tram carrying, unit person;
L is line capacity in single direction, unit per/h.

The marshaling of tram is flexible, generally from 2 to 6. It should be comprehensive considered that traffic demand and the service level of operation to choose the appropriate marshaling. The passenger capacity of tram is usually from 200 to 500 [1].

There are many factors that affect the smallest departure intervals of tram, such as vehicle control and output signal, intersection signal processing, tram braking performance, station parking time, and the tram turn back time [2]. Theoretically, in the design and operation of good modern traffic system, intersection and exhumation should not become the bottleneck of the smallest departure intervals. In a flat mouth again, the departure intervals of tram is close to 2 min in theory. The calculated tram theoretical capacity range is 6000 – 15,000 persons per hour. Because of intersection delay and stop time delay and so on, the actual capacity is always less than 10,000 persons per hour.

Due to the limitation of road conditions, the length of BRT vehicles is generally less than 20 m. Its theoretical capacity is 220 – 300 persons. The smallest departure interval of BRT is affected by intersection signal processing and station parking

time. If intersection priority completely, the smallest departure interval should be only connected with station parking time and safety driving distance, then the minimum departure intervals can reach 60 s in theory. In practical application, intersection priority completely and too small departure interval can lead to other phase in signal control motor vehicle delay greatly reducing the efficiency of the whole transportation system. Therefore, generally in practical applications, within one intersection signal cycle, it should not have two BRT vehicles reaching the intersection; otherwise, there will be a signal priority for BRT, which causes other vehicle delay greatly to increase, or two BRT vehicles line up in the intersection at the same time. The stability of the operation will be unfavorable. Therefore, the actual minimum departure intervals of BRT system is generally determined by the maximum signal cycle, and usually, set this value to 90 s [3]. The calculated BRT actual capacity range is 8800 – 12,000 persons per hour, close to tram.

Velocity

Speed reflects the quickness of public transportation. How long does it cost to reach its destination is the most concern of the passengers. So the average travel speed directly affects the applicability and popularity of the public transportation. The calculation formula of the average speed is shown below.

$$V = S/T$$

where

S is path length (m) and
T is total travel time, which includes vehicle driving time, accelerate moving time, slowdown time, intersection delay time, and station dwell.

The main factors affecting station dwell and intersection delay include average stop spacing, passenger flow volume, the door number, and intersection priority, which has a great influence on the speed of tram. Due to the differences of the above factors, the average speed of trams on the world ranges between 15 and 30 km/h. The main factors influencing BRT system also include intersection priority, average stop spacing, passenger flow volume, and the door number. On the below, we contrast these factors one by one of tram and BRT.

Maximum Operating Speed

Tram vehicle running speed can reach 70 km/h, and BRT system is 90 km/h. However, because of the limitation of station spacing and the safety point of view,

vehicle speed limit measures are taken in a straight line. Urban road transit speed is largely within 70 km/h; therefore, the affection of the both highest speed to travel time is almost the same.

Acceleration and Deceleration

Due to the traction power of the unit weight of tram and BRT close to, their acceleration is almost 1.2 m/s^2. Because the coefficient of friction of rubber tires is higher than the wheel/rail system, the deceleration of BRT is greater than tram. However, considering the comfort, we need to take limitation of deceleration, which is less than 1.4 m/s^2. Therefore, the affection of acceleration and deceleration on average speed is basically the same for tram and BRT.

Intersection Priority

Both tram and BRT can take the same form of economic benefit road right priority and signal priority. In the same line condition and signal priority measures on intersection, both the road speed limit and intersection delay time are consistent.

Stop Time on Station

If the line conditions are the same, the average station spacing and passenger flows of BRT are the same as tram. In this premise, the station stop time is mainly determined by the number of on–off passengers and the door number. Under the condition of the same capacity, the departure interval of tram is larger than BRT because of its large capacity, so tram is getting more passengers on station. Then, the door number of tram is more, so the ratio of the number of passengers getting on–off and the door number is more close to each other. Therefore, the dock station time impacting on the average travel speed is also very close.

It is discovered through comparative analysis that the traveling speed of tram and BRT is close to each other if the line conditions are the same. If the line condition is enclosed way and grade separation mode, the standards of the system has been higher than that of trams, close to the level of the light rail, so its running speed can be an average of more than 30 km/h [3]. Because the capacity and speed of the light-rail are one level higher than tram and BRT, the competition of the light-rail with tram or BRT is less. Because of the similar function and average speed, tram is fierce competition with BRT. Transportation planners and decision makers must choose one between them when selecting bus rapid transit way.

Economical Efficiency

Urban public transportation system is usually funded by the government finance as the main body, and the large and medium capacity of bus system is expensive. So, the managers of the city must consider the cost of the project when they make policy of public transportation. Unit social cost of tram and BRT is shown in Table 2.

The price of single tram vehicle is 7 – 8 times of BRT, but the tram carrying is larger and the service life is far more than BRT vehicles. Therefore, in the premise of the same capacity, the total purchase expense of tram is only 1.5 – 2 times of BRT vehicles in 30 years [4].

Because the line condition is different, the comprehensive cost of BRT and tram changes with a wide range in the world. In USA, the comprehensive cost of tram system is generally more than 3 times of BRT system, whose comprehensive cost is about $5 million/km. It has substantial agreement with European cities.

In China, for different levels of service, different vehicle configuration number, and the uneven infrastructure condition of line, the construction of tram costs about ¥70 – 100 million/km in the form of the ground and ¥100 – 200 million/km in the form of elevated, when BRT is ¥40 – 80 million/km. The reconstruction project of old road for trams is larger, when BRT is only needed to slightly adjust the old road. It is obvious that the cost of BRT is lower than tram. However, in recent years, at the background of energy conservation and environmental protection advocated, the domestic railway vehicle manufacturers have been stepping up efforts to independent innovation and technology introduction, developing its tram technology and reducing the cost. With batch production of tram in the future and the localization rate constantly improving, the average cost of tram is expected to be further reduced [5].

The average cost of tram and BRT is shown in Table 3, which is as follows.

Operating costs include personnel wages, material cost, fuel cost, and equipment maintenance cost. We usually take the unit operating costs as indicator. This chapter will take costs per passenger kilometer as comparative objects. All kinds of American public transport costs are shown in Table 4. A single travel of passenger carried by tram costs $2.31, which is less than $2.40 of BRT.

Table 2 Unit social cost of tram and BRT (Yuan per person per kilometer)

	BRT	Tram
Own cost	0.144	0.15
Road cost	0.01	0.01
Parking cost	0.00736	0.008
Traffic accident loss cost	0.0012	0
Environmental cost	0.121	0.009
Time cost	0.867	0.785
Total	1.151	0.962

Table 3 The average cost of tram and BRT

Parameters	Tram	BRT
Length (m)	36	15 – 20
Service life (years)	30	10
Cost (millions)	20	2

Table 4 Public transport costs in USA [6]

Parameters	BRT	Tram	Metro
Passenger person kilometers (million·km)	35,149	2305	21,988
Trip frequency (million of times)	5868	337	2688
Average trio distance (km)	5.95	6.76	8.21
Operating cost (million dollars)	14,066	778	4268
Operating cost per passenger kilometer (dollar)	0.40	0.34	0.19
Cost of a single travel per passenger (dollar)	2.40	2.31	1.59

The unit operating costs of BRT are slightly higher than tram, but the difference between them is small. The unit operating costs of big traffic volume of urban rail transit system are significantly lower than tram. It is reduced to 44% costs per passenger kilometer and 31% costs of a single travel per passenger [7].

The operating costs above are based on the actual passenger traffic. In the process of analysis, we need to study the trend of unit operating costs as passenger traffic. With the increase of passenger traffic, the unit operating costs of tram and BRT show a trend of decrease. Under the condition of small passenger flow, the operating cost of BRT is slightly lower than tram. And in the case of large passenger traffic, the operating cost of BRT is significantly higher than tram. The main cause leads to these differences is that in the case of passenger volume small, in order to guarantee the service level, tram could not increase the departure interval, which causes low load factor and waste capacity. So the unit operating costs increase. The capacity of BRT is less. It can extend appropriately departure intervals to improve the load factor when small passenger traffic.

The operating costs of urban public transport are mainly labor cost and energy consumption costs [8]. The capacity of tram is apparently higher than BRT, so the number of vehicle needed is less under the premise of the same passenger traffic. The driving and maintenance personnel are less, and the labor cost is decreased. Furthermore, the energy consumption of tram is significantly less than bus of tire-running. The unit operating costs of tram are less than BRT in conclusion.

Table 5 Parameters of vehicle

Parameters	BRT	Tram
Size (m)	12 × 2.5 × 3.15	45 × 2.4 × 3.4
Dead weight (t)	11.3	53.2
Rated passenger capacity (person)	80	288

Table 6 Result of unit energy consumption (joule per person)

Load factor (%)	Velocity											
	BRT						Tram					
	20	30	40	50	60	80	20	30	40	50	60	80
50	179	303	462	669	934	1676	45	74	107	147	194	315
75	131	220	334	479	664	1179	33	53	77	106	140	228
100	107	179	269	385	530	930	26	43	62	86	113	184
125	93	154	231	328	449	781	23	37	54	74	97	158

Energy Consumption

Encouraging public transit and vigorously developing public transportation are the main way to reduce energy consumption per capita travel and also are the main way to sustainable development of the traffic, so the unit energy consumption is an important comparative index of tram and BRT.

The unit energy consumption is defined as the energy dissipations of vehicles carrying one passenger per kilometer. The parameters of vehicle set up are shown in Table 5.

The unit energy consumption variation trend with load factor and velocity is calculated as shown in Table 6. It is obvious that the theoretically unit energy consumption of tram is much lower than BRT. On the premise of full load, the unit energy consumption of BRT is 4 – 5 times of tram.

Influence of Speed on the Unit Energy Consumption

On the premise of certain load factors, speed is the only factor affecting the unit energy consumption. Figure 1 reflects the variation trend of the ratio of unit energy consumption of BRT and tram as speed on the case of full load. With the increase of speed, the unit energy consumption of BRT increase amplitude is distinctly larger than tram. The more the speed, the greater the energy-saving advantages of tram.
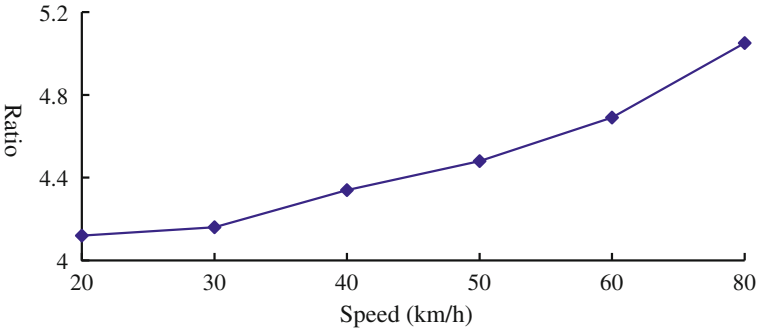
Fig. 1 The variation trend of unit energy consumption as speed
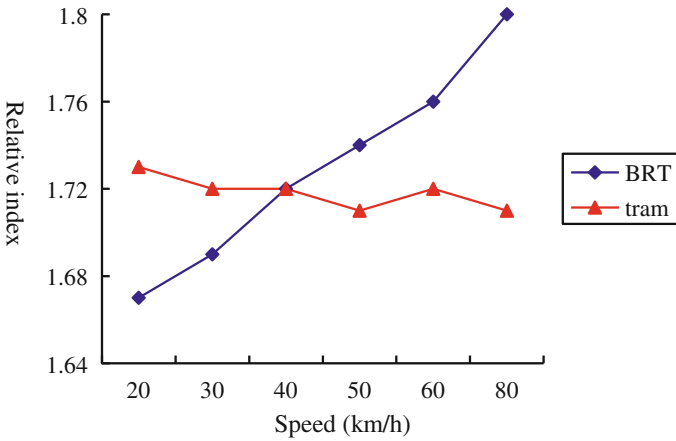
Fig. 2 Relative energy consumption index of 50% load factors

Influence of Load Factor on the Unit Energy Consumption

In order to reflect the variation trend of energy consumption under different load factors, we take the ration of energy consumption under different load factors with full load as the relative energy consumption index, as a comparison and analysis of reference quantity. It is generally accepted that the influence of load factor on the unit energy consumption of rail transport is greater than the bus, because of the heavier rail transit vehicles. When the speed is less than 40 km/h, it has a little influence in load factor on unit energy consumption to tram and BRT. As the speed exceeds 40 km/h, the energy consumption of BRT increases continuously expand in case of 50% load factors, more than tram. The unit energy consumption of tram is greatly influenced by the load factors with low speed. As shown in Fig. 3, the higher the speed, the greater the influence of load factors on energy consumption to BRT (Fig 2).
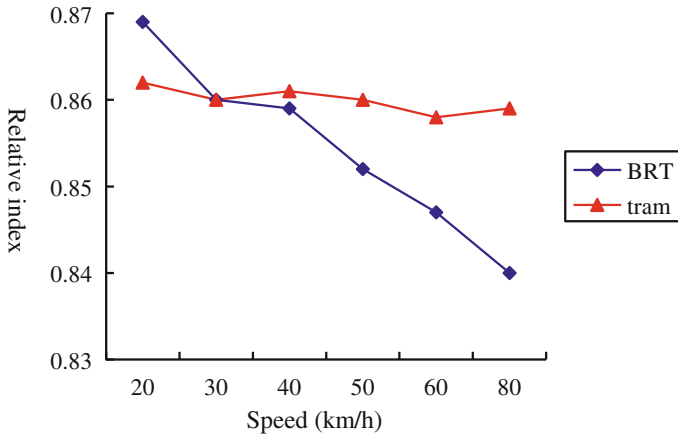
Fig. 3 Relative energy consumption index of 125% load factors

Conclusion

To medium capacity of BRT and tram, this paper comparative analysis respectively from the aspects of capacity, speed, economical efficiency, energy consumption. In conclusion, the average speed and capacity of both are almost, and the economical efficiency of BRT is little better than tram, while the energy consumption of tram is greater than BRT.

In addition, the trams' green environmental protection and no pollution is good for the image of cities. Tram was regarded as a form of city environmental protection and traffic model of mutual integration. The capacity of 10,000 passengers per hour lines is close to the limits of BRT, whereas tram could increase carriages through the marshaling form, increasing its capacity and providing the highest cost-effective transport service. On Comprehensive consideration, it is suggested that policy makers give priority to tram under the same condition.

References

1. Wei, Chao. 2008. Study on the applicability of modern tram. Tongji University Master Degree Thesis.
2. Transit Capacity and Quality of Service Manual. 2003. 2nd ed. (TCRP Report 100). Transportation Research Board of the National Academies.
3. Wang, Yancai. 2011. The preliminary research to application of modern tram. Nanjing Forestry University Master Degree Thesis.
4. Zhang, Huajin. 2013. Comparison of modern tramways with BRT and application in China. *Rail Transit*.
5. Long, Qianqian, Haibin Ren, and Xicheng Ma. 2014. Comparison and analysis of modern tram with BRT. *Electric Locomotives & Mass Transit Vehicles*.

6. Evans, John (Jay). 2005. Capacity and Cost Comparisons of Rapid Transit Modes. *Institute of Transportation Engineers Annual Meeting*, 8.
7. United States General Accounting Office. 2001. Report to Congressional Requesters. MASS TRANSIT-BRT Shows promise, 9.
8. Mao, Baohua. 1999. *Operation management of urban rail transit system*. China: China Railway Publishing House.
9. Huo, Kaiyu. 2014. Study of the construction feasibility and technical proposal of modern tram. Southwest Jiaotong University Master Degree Thesis.
10. Cui, Yanan. 2012. Study of the application mode and regional applicability evaluation of modern trams. Beijing Jiaotong University Master Degree Thesis.

Analysis of Speed Characteristics of Different Types of Drivers at a Certain Cruise Speed

Yan Xing, Lishuang Sun, Lianghong Ji and Weidong Liu

Abstract In order to analyze the speed characteristics of different types of drivers in the case of cruise at a certain speed, 50 drivers were tested on this condition. First of all, using clustering analysis method, the drivers were divided into 5 categories according to the different driving styles. Then, analyze the mean and standard deviation of the relative velocity of the 5 kinds of drivers at the speed of 60/90/120 km/h. Obtain the relative velocity distribution characteristics of different types of drivers in different speed conditions. Through this study, it found that the skilled drivers could control the vehicle speed more accurately than the non-skilled drivers; the robust drivers were more frequent than the aggressive drivers to adjust speed. When the road type changed, the driver was usually difficult to accurately control the speed of the vehicle.

Keywords Cruise control · Cluster analysis · Instantaneous velocity · Relative velocity

Introduction

Cruise Control System (CCS) can make the cars work in the favorable range of the engine speed, reduce the driver's labor intensity in driving operation, and improve ride comfort. At present, it has been widely used in assistance technology for vehicles. In recent years, many experts and scholars have carried out relevant research in this field. Donn H. Mcmahon et al. in 1989 established the vehicle longitudinal motion model,

Y. Xing (✉)
School of Automotive Engineering, Jilin University,
Changchun 130022, Jilin, China
e-mail: 44050412@163.com

Y. Xing · L. Sun · L. Ji · W. Liu (✉)
School of Transportation Engineering, Shenyang Jianzhu University,
Shenyang 110168, Liaoning, China
e-mail: 359063088@qq.com

which comprehensively considered the vehicle's longitudinal movement involved many factors and the precision was higher. But it also led to the complexity of the model relatively higher, the amount of computation was very large, and it was difficult to carry out practical application [1]. P. Ioannou et al. was based on the longitudinal motion model of multiple state vehicle in 1993 [2]. Z. Xu et al. carried out the linearization process based on the adaptive control of the reference model and gave the control rate based on continuous time and discrete time [3]. The controller designed by M.K. Liubakla et al. is adaptive control based on gradient algorithm, which can make the control algorithm more convenient [4]. Zhan Jun et al., according to the operation characteristics of cruise control, introduced a two-state model of engine and gave the driving and braking switch standard [5]. Hou Dezao et al. established the longitudinal motion model of the vehicle according to the constant speed control for vehicle longitudinal model requirements [6]. Qu Zhaowei, Xing Yan, and others by analyzing the discrete characteristics of the vehicle in the intersection determined the relationship between driver's driving psychology and motor vehicle running characteristics [7–9]. The research findings were mainly on the technical performance of cruise control and only few studies on characteristics of drivers under the cruise. The study of speed variation in cruise condition is fewer, based on the characteristics of different types of drivers. However, in the actual situation, the change of the speed of the motor vehicle directly reflects the driving comfort and safety of the driver in the case of cruise at a certain speed. 50 drivers were tested in the case of cruise at a certain speed. First of all, using cluster analysis method, the drivers were divided into 5 categories in accordance with the different driving styles. Then, analyze the mean and standard deviation of the relative velocity of the 5 kinds of drivers at the speed of 60/90/120 km/h. And obtain the relative velocity distribution characteristics of different types of drivers in different speed conditions.

Driver Sample Data Processing Method

In order to meet the needs of the project, 50 healthy people with a driver's license were selected as the sample of the project. The samples contained not only professional drivers whose driving experience was more than 10 years, but also novices whose driving experience was not over 3 years. The selected testers included different age, sex, driving mileage in order to make the test data more accurate. In the 50 drivers, male drivers were 38, occupied 76% of the total number. Female drivers were 12 people, occupied 24% of the total number.

Driver Sample Clustering Process

In order to analyze the driving behavior conveniently, cluster 50 drivers. The classification indexes of cruise condition are acceleration and relative velocity.

The indexes of the following conditions are the relative distance and collision time (TTC). Because the sampling points are the points obtained in a period of time, the processing of each indicator of the data is to obtain the average value and standard deviation. For each driver sample, its mode Z could be represented in formula (1):

$$Zk = (F1k, F2k, F3k, F4k, F5k, F6k, F7k, F8k), \quad k = 1, 2, \ldots 50 \quad (1)$$

where $F1$ is the mean value of acceleration, $F2$ is the acceleration variance, $F3$ is the average of the relative velocity, $F4$ is the standard deviation of relative speed, $F5$ is the average of the relative distance, $F6$ is the standard deviation of the relative distance, $F7$ is the average of the collision time, and $F8$ is the variance of the collision time. In the experiment, the constant speed cruise mode was divided into straight constant cruise control, curve constant speed cruise, ramp constantcruise control, and several tests were performed respectively on 60km/h, 90km/h and 120km/h. In the process of data processing, in order to not lose the general, fuse the acceleration data of different speeds and the relative speed data both from a same driver. In the following conditions, use the same way to deal with relative speed and relative distance. In clustering process, Euclidean distance was used as the distance measured between models to describe the similarity between patterns. Take two driver patterns $Zi = (F1i, F2i, F3i, F4i)$ and $Zj = (F1j, F2j, F3j, F4j)$; Euclidean distance $d(Zi, Zj)$ is shown in formula (2):

$$
\begin{aligned}
d(Zi, Zj) &= ||Zi - Zj|| \\
&= [(F1i - F1j)2 + (F2i - F2j)2 + (F3i - F3j)2 + (F4i - F4j)2]1/2
\end{aligned}
\quad (2)
$$

Because the drivers' status is complicated and changeable, there is no clear boundary between the man and man. It is inappropriate to divide drivers into a certain class by the traditional hard classification method. Fuzzy mathematics as a kind of effective method to describe the uncertainty of things in the field of artificial intelligence is suitable for the classification of the drivers, so the data processing is based on the fuzzy mathematics method of cluster analysis. The quality of clustering is closely related to the number of clustering. The clustering number is too complex to explain and analyze the clustering results. However, if the clustering number is too small, it will cause the loss of information and mislead the final decision. Therefore, the optimal clustering number Mo is determined by several clustering validity indexes. Selection of indicators includes partition coefficients (PC), Xie-Beni index (XB), partition index (SC), and Dunn index (DI). These indicators represent the interclass aggregation of clustering and interclass separability. The smaller the index value, the better the clustering effect. The definition of each index can refer to the literature. Figure 1 shows the variation of effective index with clustering number M. Drawn from the graph, Xie-Beni index (XB) and Dunn index (DI) both reach the local minimum value, when the clustering number M is 5. Partition coefficients PC and partition index SC gradually descend with the increase of M. The variation range of M = 5–9 is very small, and the optimal clustering

number is $Mo = 5$ after the comprehensive consideration. The specific clustering validity indexes change with the number of clustering. As shown in Fig. 1.

Based on the optimal number of cluster number $Mo = 5$, the driver's driving behavior characteristic is obtained. As shown in Fig. 2:
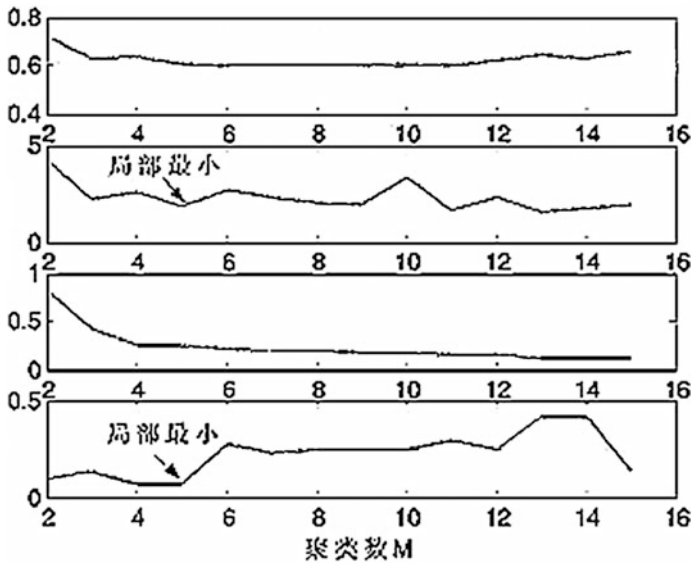
Fig. 1 The change of clustering validity indexes with the number of clustering
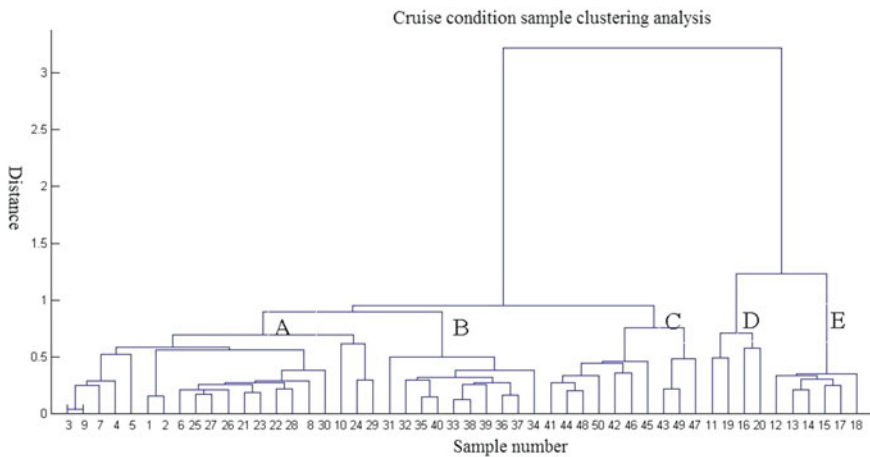
Fig. 2 Clustering pedigree chart of the sample

This clustering result is acceptable. The five clustering represents the driving style of the five drivers. Each category of driver's driving style is described as shown in Table 1.

Secondly, the characteristics of different types of driver behavior were analyzed.

Least Squares Fitting of Statistical Parameters

In order to get the relation curve between the mean value of car following distance and the velocity, the relation curve between the mean value of time interval and the velocity, the least square fitting is used to fit the original data.

For the given data points $\{(Xi, Yi)\}(i = 0, 1, \&\&, m)$, in the class of the function Φ, $p(x) \in \Phi$, make the sum of square $E2$ of errors the smallest, $E2 = \sum [p(X_i) - Y_i]^2$. In terms of geometry, it is to seek curve $y = p(x)$, which is the minimum sum of square of the distance between the given points $\{(Xi, Yi)\}(i = 0, 1, \&, m)$. Taking the time interval under the condition of steady state for example, a class driver's original data scatter diagram and fitting out the graph are shown in Fig. 3.

In this paper, if not specified, the curves given in this paper are fitted curves.

Category	Sample number	Feature description
A	20	Proficient and steady
B	10	Skilled and aggressive
C	10	Neutral
D	4	Unskilled and not aggressive
E	6	Unskilled and not steady

Table 1 The number of cluster samples and the description of the driver's driving behavior

(a)

Theoriginal datascatter diagram

(b)
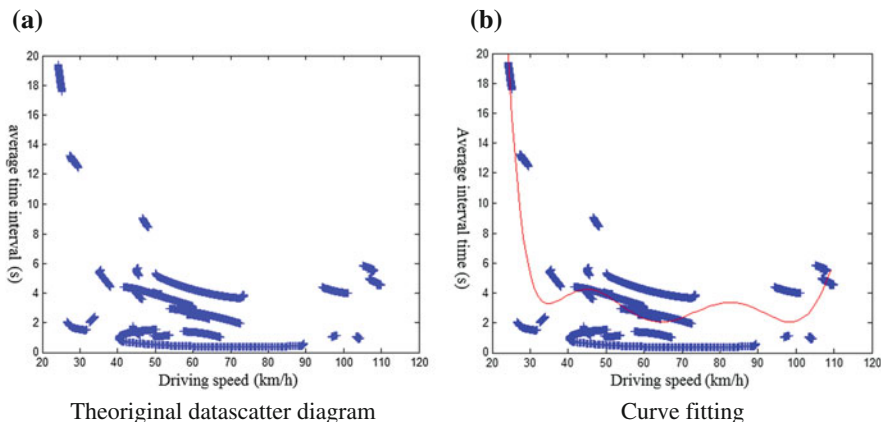
Curve fitting

Fig. 3 The least-squares curve fitting

Analysis of Driving Speed Change in the Cruise Conditions

Cruise control is a part of the ACC function, when there is no guide car, and ACC is in the open state, it means to enter the cruise state. According to the road environment and the operation of the drivers, cruise can be subdivided into straight cruise, bend cruise, ramp cruise, etc. The relative speed defined in this paper is the difference between the speed Vhost(t) and the setting speed Vset in the cruise condition, which is shown in formula (3):

$$Vr(t) = Vhost(t) - Vset \qquad (3)$$

A statistical description of the relative speed of different types of drivers in the driving process is shown in Tables 2, 3, 4, 5, and 6:

From Tables 2, 3, 4, 5 and 6, it can be seen that the relative speed of the skilled drivers is smaller than that of the unskilled that means to control the speed of the vehicles more accurately under the conditions of different road and different speeds. The standard deviation will be even greater, when the robust drivers are compared with the aggressive drivers. The most direct manifestation is that the speed fluctuation amplitude is bigger or the speed adjustment is more frequent in the course of moving vehicles. Meanwhile, through the comparison, it is found that the relative speed standard deviation of the steady drivers has a tendency to decrease with the increase of speed, but the aggressive drivers' change is not obvious.

Table 2 Statistic description of driving speed of Class A drivers

Speed road type	60 km/h		90 km/h		120 km/h	
	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation
Straight line	2.9685	3.0004	−2.6784	1.9211	−0.4822	0.4050
Curve	3.1851	0.9008	−0.1268	2.0591	3.4724	0.0703
Ramp	−0.7595	2.2615	0.4585	3.1515	−6.1027	1.0909
Large flow density	1.8517	2.0248	−2.1616	2.4595	−4.5795	4.1874

Table 3 Statistic description of driving speed of Class B drivers

Speed road type	60 km/h		90 km/h		120 km/h	
	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation
Straight line	−4.9485	3.7147	−1.2792	2.5706	−3.1182	1.6697
Curve	−3.3972	1.5889	−2.8177	1.3663	2.0632	3.1414
Ramp	−3.9991	4.3024	−3.7238	2.9836	1.0864	3.6382
Large flow density	−1.4939	2.7057	−1.1473	3.4624	0.1631	4.7426

Table 4 Statistic description of driving speed of Class C drivers

Speed road type	60 km/h		90 km/h		120 km/h	
	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation
Straight line	−1.7216	3.8660	−0.2469	2.5094	−1.5779	3.4082
Curve	0.8231	2.7261	−0.5917	2.2928	−0.1340	2.6625
Ramp	−2.1256	2.3242	−1.6142	4.4991	−2.0400	4.3292
Large flow density	−0.9728	3.1236	−0.6030	2.5100	−2.5389	1.8780

Table 5 Statistic description of driving speed of Class D drivers

Speed road type	60 km/h		90 km/h		120 km/h	
	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation
Straight line	−8.0845	4.8790	0.2403	2.7817	−4.9371	2.1213
Curve	−2.4121	2.2741	−5.9731	6.6832	4.3854	1.0491
Ramp	1.0766	4.2809	1.5981	5.7742	2.5532	6.0708
Large flow density	−1.0510	4.7323	−2.8433	4.1456	−4.2478	1.8981

Table 6 Statistic description of driving speed of Class E drivers

Speed road type	60 km/h		90 km/h		120 km/h	
	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation	Mean value (km/h)	Standard deviation
Straight line	9.7821	4.7346	−1.5149	0.7915	2.2833	2.4722
Curve	1.5070	3.0197	8.0044	1.3652	12.9533	1.3628
Ramp	2.5656	2.8640	−0.9141	5.0432	−4.3925	2.6504
Large flow density	3.9966	4.0307	7.6223	1.7903	−0.7092	1.0293
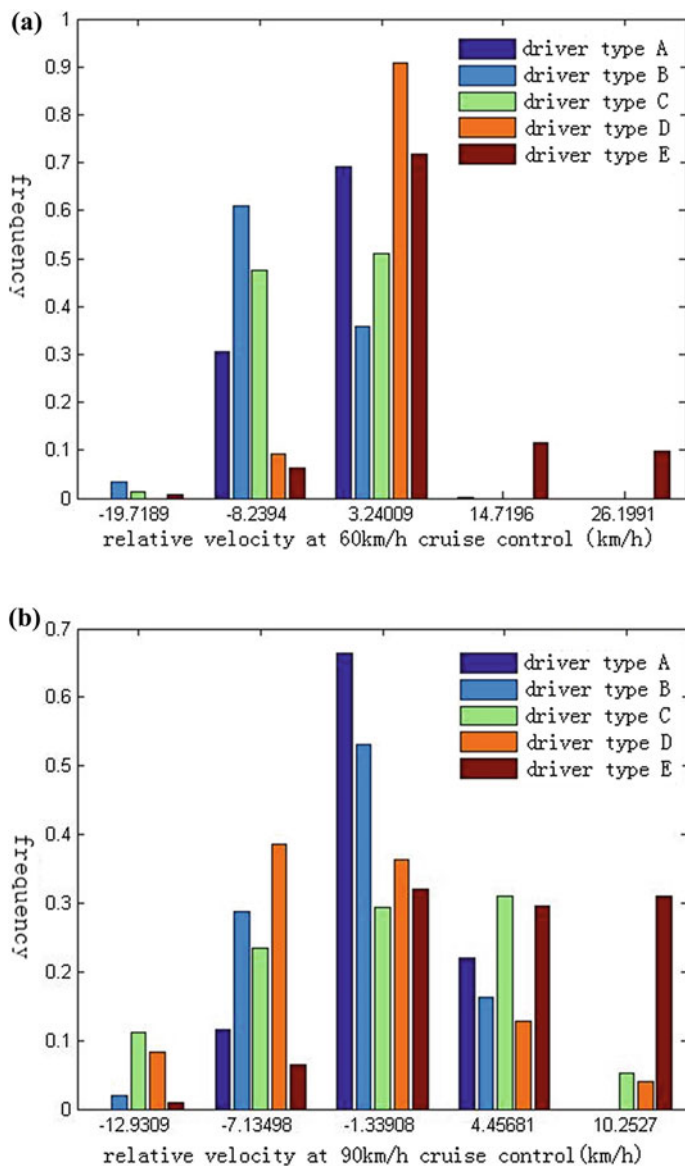
Fig. 4 Relative velocity frequency distribution of different types of drivers under different speeds
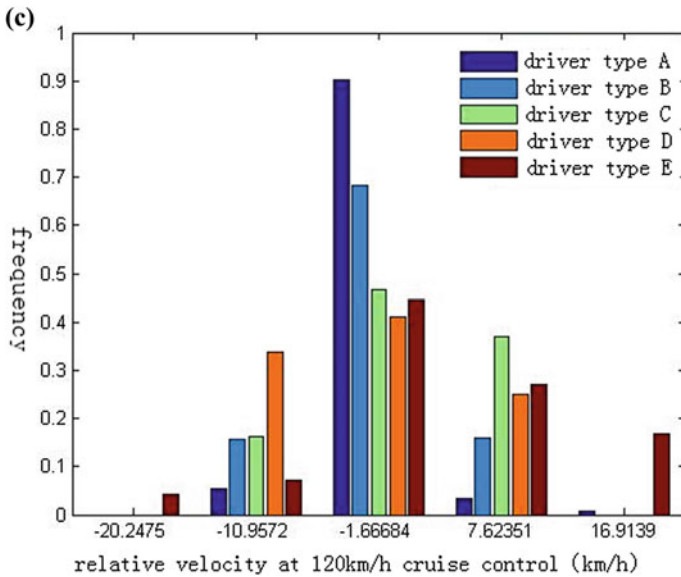
Fig. 4 (continued)

Analysis of Velocity Frequency Distribution of Different Types of Drivers in the Cruise Conditions

Carry on statistics to the frequency distribution of relative velocity of five kinds of drivers A, B, C, D, and E under different speeds. Statistical graphs are shown in Fig. 4.

What can be seen from the graph is that the Class A drivers' relative velocity distribution is more concentrated at different speeds, and there is little deviation with the set speed. Performance characteristics of the drivers are in accordance with the clustering; compared with A, although the relative speed of B is also mainly distributed in about 0, B drivers' relative velocity distribution is more widely distributed than A; Class C drivers on the basis of the Class B are the more "short" and more "fat"; in the five categories of drivers, it is neutral. The distribution of Class D and Class E is not obvious compared with A and B.

Analysis of Relative Velocity Distribution of Different Types of Drivers in the Cruise Conditions

Extract the data from skilled and unskilled drivers in the sample and obtain the relative velocity distribution in the whole velocity range. As shown in Fig. 5.
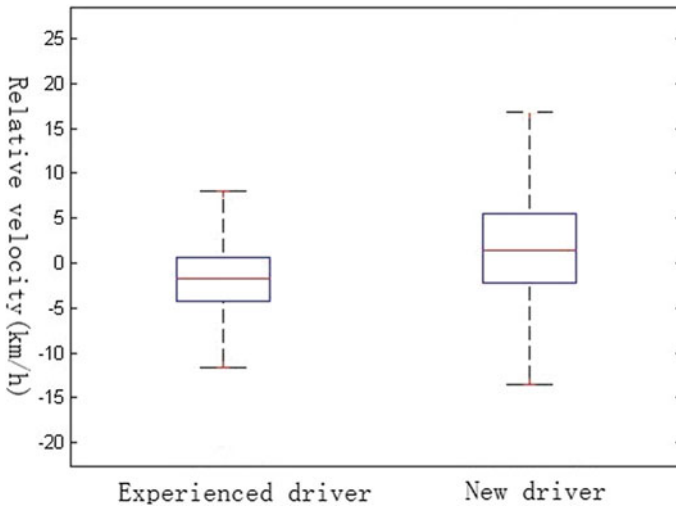
Fig. 5 Relative velocity frequency distribution of different types of drivers under different speeds
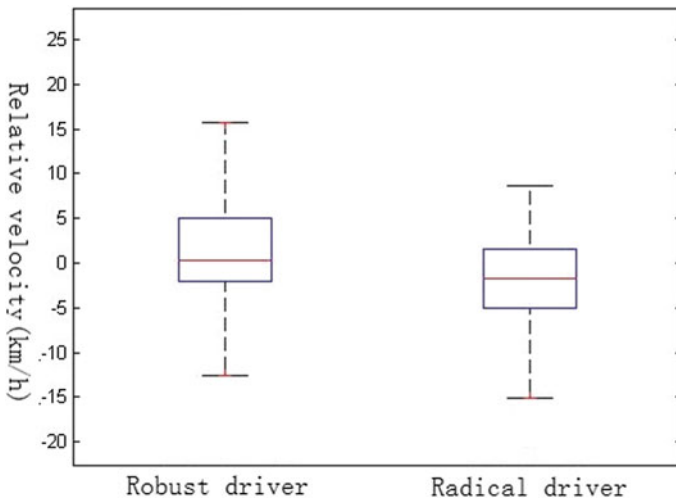
Fig. 6 Relative velocity frequency distribution of different types of drivers under different speeds

What can be seen from the graph 5 is that the skilled drivers are more accurate in the control of the speed and the relative velocity distribution is smaller.

Extract the data from robust drivers and aggressive drivers in the sample and obtain the relative velocity distribution in the whole velocity range. As shown in Fig. 6.
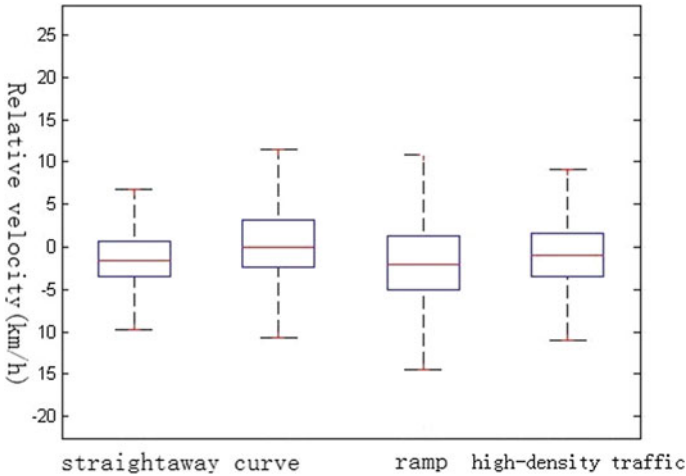
Fig. 7 Relative velocity statistics of the average drivers in different road types

The range of relative velocity distribution of robust drivers is slightly wider than the aggressive drivers, but the median relative velocity is more close to 0.

In order to get the difference of the average driver's control behavior in different road conditions, after removing the singular value and the extreme value, count the relative velocity distribution of the five kinds of drivers in the whole speed range. As shown in Fig. 7.

What can be seen from the graph 7 is that the distribution of the relative speed of the driver is obviously different along with the change of the road type or the traffic condition. For cruise control condition, the driver has to control the vehicle at the specified speed driving; therefore, the mean value and variance of relative speed can be used to judge whether the driver controls the vehicle accurately or not. When the road type changes, we can find that the distribution of relative velocity is more and more wide. Mainly in the corners, the driver needs to correct the steering wheel angle to maintain the correct attitude of the vehicle and cannot adjust the speed in time. In the case of the ramp, whether uphill or downhill, both have a greater impact on the speed of the vehicles. If the drivers fail to adjust the throttle opening, it will affect the relative speed; under the large flow density of the straight channel, the relative velocity amplitude is increased.

Summary

In this paper, the driving characteristics of different types of drivers under different cruise speeds were analyzed by collecting the parameters of vehicle speed, relative speed and other parameters. The results show that the skilled drivers can control the vehicle speed more accurately than the non-skilled drivers; the robust drivers are

more frequent than the aggressive drivers to adjust speed. When the road type changes, the driver is usually difficult to accurately control the running direction and the speed of vehicles. With the great increase of the traffic flow density, the fluctuation range of the motor vehicle speed increases gradually.

References

1. Donn H. McMahon, J.K. Hedrick. 1989. Longitudinal Model Development for Automated Roadway Vehicles. California PATH research report, UCB-ITS-PRR-89-5.
2. P. Ioannou, Z. Xu, S. Eekert, D. Clemens, T. Sieja. 1993. Intelligent Cruise Control: Theory and Experiment. In *Proceedings of the 32nd Conference on Decision and Control* , 832–845.
3. Z. Xu, P. Ioannou. 1994. Adaptive Throttle Control for Speed Tracking. University of Southern California PATH Research Paper, UCB-ITS-PRR-94-09.
4. M.K. Liubakka, D. S. Rhode, J. R. Winkelman, P.V.Kokotovic. 1993. Adaptive Automotive Speed Control. IEEE TRANSACTIONS ON AUTOMATIC CONTROL, 599–612.
5. J. Zhan. 2006. Setup of vehicle longitudinal dynamic model for adaptive cruise control. *Journal of Jilin University* 2006 (02): 157–160.
6. D. Hou, F Gao, L. I. Keqiang. Vehicle longitudinal dynamic model for vehicle collision avoidance systems. *Journal of Tsinghua University* 2004(02), 258–261.
7. Zhaowei Qu, Xing Yan, 2013. The departure characteristics of traffic flow at the signalized intersection. *Mathematical Problems in Engineering*. 2013(1), 1–11.
8. Zhaowei Qu, Xing Yan, 2012. A Study on the Coordination of Urban Traffic Control and Traffic Assignment. Discrete Dynamics in Nature and Society, 2012 (1). 1-12.
9. Yan, Zhaowei Qu, 2016. Study on Vehicle Delay Based on the Vehicle Arriving Distribution at Entrance Lanes of Intersection. *Procedia Engineering* 2016 (137) : 599–608.

A Research on Traffic Conflicts Between Vehicle and Pedestrian on Urban Typical Road Section

Qian Cheng, Leyi Wang, Chenggang Li, Xiaobei Jiang and Wuhong Wang

Abstract In a typical road traffic environment, drivers may make wrong judgments and predictions to the pedestrians ahead because of a variety of reasons. Real pedestrian-traffic conflict data are collected and analyzed. This research videotaped conflicts between pedestrians and vehicles in three non-signal pedestrian crosswalks in Beijing. A conflict safety distance model of typical road sections is proposed by identifying the process of the changing of vehicle's speed and pedestrian's speed. Reasonable suggestions for the development of the vehicle driving assistant systems are presented.

Keywords Traffic conflict · Vehicle safety driving characteristics · Pedestrian crossing behavior characteristics · Safe distance model

Introduction

With the increase of the number of vehicles, more and more traffic accidents have happened constantly, causing high traffic casualties and economic losses [1]. Traffic accidents occurred for many reasons, which can be divided into vehicle factors, driver factors, traffic environmental factors, pedestrian factors, etc. [2]. Data shows that among all factors 70–80% traffic accidents were caused by the driver factor [3]. Whether the driving behavior is accurate and reliable directly determines the safety of the road traffic. Therefore, it is very important to study the driving behavior of the driver.

A traffic conflict can be expressed as an effective contingency measure (deceleration, change direction) to avoid traffic accidents, which is not actually an acci-

Q. Cheng · L. Wang · C. Li · X. Jiang (✉) · W. Wang (✉)
Department of Transportation Engineering,
Beijing Institute of Technology, Beijing 100081, China
e-mail: jiangxiaobei@bit.edu.cn

W. Wang
e-mail: wangwuhong@bit.edu.cn

dent but has a conversion relationship with accident [4–6]. Considering the current situation of traffic in China, this paper chooses a typical and common form of conflict. Then a multidimensional analysis under typical road section traffic conflict situation to the traffic behavior characteristic is proposed.

In this paper, conflicts between vehicles and pedestrians crossing streets are selected to explore the hedging measures taken by drivers and pedestrians in face of conflict. A calculation model of vehicle transverse and longitudinal safety displacement based on driving perception and the definition of safety zone are proposed. This paper analyzes the concept of vehicle driving safety region model and its deterministic form, and analyzes the behavior of motor vehicles and pedestrians in this conflict situation. Based on this, this paper puts forward some suggestions to reduce such conflict, and based on the acquired conflict data on the development of driving support system makes recommendations

Method

Conflicts Types

In order to obtain detailed and accurate conflict data in this survey, the selected conflict location should be a typical road section of the city, select one-way single lane or two-lane road section, with a certain traffic flow and pedestrian flow (Fig. 1). Three traffic sections, such as South Road, Haidian District, Beijing, were chosen as the survey sites to observe the traffic conflict.

According to the arrangement of traffic conflict observation time, the period of traffic conflict observation on the same location is 2 days or 3 days. The observation time is 7:00–8:00, 11:00–12:00, and 5:00–7:00 every day for the observation of the pedestrian peak period to get more conflicting samples.
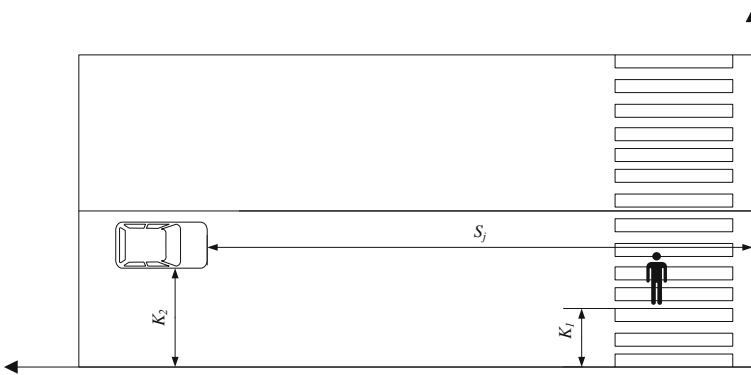
Fig. 1 Schematic diagram of conflict

Table 1 Conflict sample

Sample 1 (0.5 s)	Male	Volkswagen Polo	Weigongcun Street
Veh location (m)	Veh speed (km/h)	Ped location (m)	Ped speed (m/s)
48.922	34.836	0	0.757
45.02	28.097	0.378	0.136
41.185	27.613	0.446	1.054
37.387	27.341	0.973	1.666
33.613	27.176	1.806	0.574
29.869	26.954	2.093	0.556
26.147	26.8	2.371	0.735
22.491	26.323	2.739	0.839
18.68	27.439	3.158	0.612
14.888	27.303	3.464	0.646
10.964	28.252	3.787	0.626

Conflicts Data

Speed, types of vehicles, and pedestrians were collected from the video. One single conflict example is showed in Table 1.

Data Analysis

Analysis of Individual Driving Behavior

After obtaining the video of the crosswalks, the motor–pedestrian interaction types are classified into the following four categories:

(1) Pedestrian and vehicle observed each other, after judging that there is no potential danger, they did not change their speed. This situation is considered safety, it was not included into the conflict samples.
(2) Pedestrian and vehicle realized that there might be a possible conflict. And then vehicle decelerated, pedestrians sped up after the judgement. He walked directly to the crosswalk to avoid the occurrence of conflict, the vehicle continued to keep moving slowly. This situation is captured into the conflict samples.
(3) Pedestrian and vehicles observed each other after the form of movement and determines the possible conflict. And then the vehicle accelerated, pedestrians slowed down or even come back to the road line. After waiting for the vehicle pass through the crosswalk, pedestrian would walk through the crosswalk. This situation is captured into the conflict samples.

(4) Vehicles take emergency braking, emergency lane, and other forms to avoid traffic conflicts, this situation can easily lead to traffic accidents, and belong to the more serious traffic conflict, and it was collected into the conflict samples.

Analysis of Individual Pedestrian Crossing Behavior

In different sections, the average speed of vehicles directly affects the speed of crossing pedestrians. Maybe in faster speed sections, pedestrians feel more dangerous, which will deliberately make them to speed up the pace. Table 2 is the survey of the three sections of the average speed and the average pedestrian crossing speed of the relationship.

Analysis of Conflict Samples

The conflict severity is divided into three categories: high-risk conflict, dangerous conflict, low-risk conflict, conflict distance (the first time when the vehicle decelerates), conflict deceleration (the first time for the motor vehicle Deceleration time), the time of collision (TTC) is the clustering index when the first speed of vehicle or pedestrians changes greatly (greater than 20%).

In order to facilitate the analysis and comparison, the clustering index is normalized, and the observed values of all variables are transformed to 0–1 due to the different units of the variables and large numerical difference. The specific treatment method is to use the actual observation value divided by the maximum observed value of the sample, the results of the standardized postprocessing cluster analysis, the results are shown in Table 3.

(1) TTC

According to the clustering analysis results, we get the TTC distribution diagram, as shown in Fig. 2. Combining the TTC analysis of conflicting samples, one can estimate the critical value between the various hazard levels in this conflict situation. In Fig. 2, you can choose 1 s for the high-risk and dangerous situation of the critical value, select the 1.5 s for dangerous and low-risk situation critical value. This reference value can be used as a measure for pedestrian clearance and the development of a driving assistance system.

Table 2 Relationship between vehicle and pedestrian

Position	1	2	3
Average vehicle speed (km/h)	31	34	38
Average pedestrian speed (m/s)	1.033	1.233	1.25

Table 3 Clustering index of partial conflicts

	Conflict distance (m)	Conflict decelerate speed (m/s^2)	TTC (s)	Result
1	27.314	2.999	2.269	2
2	33.613	3.744	2.474	2
3	30.686	3.181	1.759	2
4	31.701	2.651	1.895	2
5	27.786	1.872	1.593	2
6	19.895	5.798	0.66	1
7	26.544	3.452	1.335	3
8	19.997	2.986	1.188	3
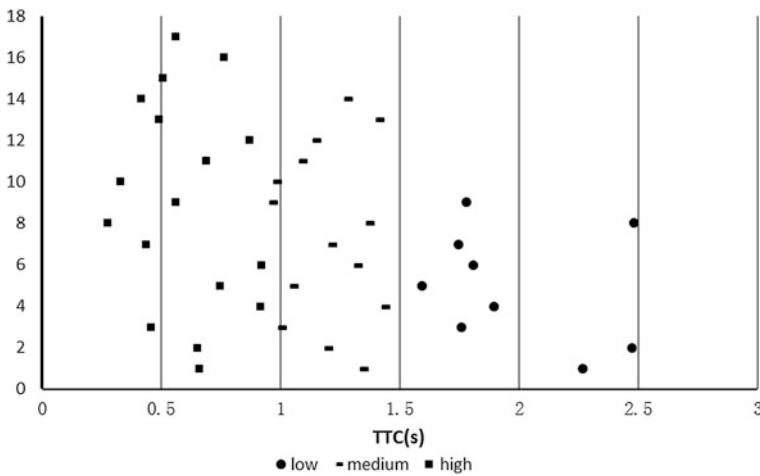
Fig. 2 TTC distribution after clustering analysis

(2) Conflict distance

From Table 4 can be drawn in the three kinds of risk levels under the mean of the conflict distance, the results of the conflict level to determine the significance of guidance. For example, if pedestrian cross-street behavior begins to take place before 32 m from the crosswalk, and take the initiative to hedge, then the probability of a traffic collision is low, even if there is a traffic conflict and its severity is low. In this case, there will be no property damage, and the efficiency of the road will not be affected.

Such values can be used in the development of a driving assistance system. For example, if the severity of the conflict is low, it can be used to "provide information" to alert the system and provide driver information such as voice reminders, whistle warnings, etc. If a conflict is identified from dangerous level, a "cooperative" approach, by assisting the driver to complete deceleration braking and other acts to avoid conflict; if the judge is a high risk of conflict, then take the

Table 4 The sample features at each level of conflict

	Amount	Mean (m)	Variance (m)
Low	9	28.997	13.310
Medium	14	23.206	8.982
High	17	15.710	5.711

"alternative" approach, by some obvious agency to operate on behalf of the driver to avoid traffic conflicts.

Result

Pedestrian Conflict Features

We make the following conclusions by analyzing of the extracted pedestrian data:

(1) In typical urban roads, the actual situation of traffic flow on the pedestrian street has a great impact on pedestrian. When the traffic flow is low density, pedestrians are more likely to find the chance to cross the street, even if the conflict occurs, the severity would to be too high. In the case of high traffic density, pedestrians are difficult to find the chance to cross the street, even if the speed is low.

(2) Pedestrian crossing is affected by the average speed of road sections. On the sections with faster average speed, pedestrians usually adopt the one-off approach to reduce the duration of potential danger, and the pedestrian average speed is higher. In the sections of the slower average speed, pedestrians crossing the crosswalk will pay more attention to the surrounding traffic environment, after passing through a lane will temporarily stay to observe the next lane to determine the actual situation, the duration of such a potential danger Longer, but the safety is not reduced.

(3) After the anticipation of possible conflict, pedestrians responded to this stimulus in two forms: one for emergency stop or return to the roadside, this is a conservative behavior, mainly in pedestrians to predict themselves the vehicle may enter the danger zone during the time period during which it passes through the danger zone. Such behavior is more common in the elderly, women, and other conservative population. The other is to accelerate through, mainly in pedestrians estimated motor vehicle deceleration, they have the ability to pass before the arrival of motor vehicles in dangerous areas, the data show that in this case, the average pedestrian speed increased by 20–30%, this is radical type of behavior, such behavior is more common in young people.

Vehicle Conflict Features

For the vehicle, the information and situation to be observed are more complex, not only to observe the situation of pedestrians ahead, but also the information of traffic signs and other motor vehicles in the lane. Therefore, when passing a typical road section, motor vehicle drivers need to pay attention to more information than pedestrians. Conclusions for vehicles are as below.

When the driver observes the presence of pedestrians, driver will firstly perform the anticipation of the conflict. Generally, the vehicle will start to decelerate at about 28.5 m (about 5 m) from the crosswalk. In the motor vehicles which adopting deceleration, the deceleration rate was positively correlated to the speed. In addition, the deceleration behavior is also affected by vehicle type, road type and other factors. Overall, however, the magnitude of the deceleration depends on the perceived risk of potential conflict by the driver. It is concluded that, in the vehicle driving support system development module, it is recommended that the vehicle identification mark be set up in front of the main crosswalks of the road section as the starting signal of the vehicle pedestrian identification system. After receiving the start signal, the vehicle's own detection device begins to identify whether there is pedestrian crossing in front.

After the first deceleration of motor vehicles, vehicle behavior is divided into two kinds: one case, motor vehicles continue to maintain low speed, pedestrians accelerate to go through the crosswalk, in this case, the motor vehicle speed in the slowdown 60–70% of the former speed fluctuations, this is a conservative driver behavior. Alternatively, the driver observes that the pedestrian stops moving forward or the speed decreases and the driver accelerates through the crosswalk, where the speed of the motor vehicle will generally return to or even exceed the speed before deceleration, which is radical driver behavior before the deceleration of the vehicle speed of 100–120%, due to the current traffic conditions in China, the lack of effective communication between drivers and pedestrians, in this case prone to traffic accidents.

To sum up, after the first slowdown, in order to reduce the loss of conflict, strengthening the communication between pedestrians and motor vehicles is needed.

Discussion

There exists two ways to induce human–machine conflict in a typical road section of city: one is that the two sides of the conflict cannot observe each other because of the limitation of line of sight or lack of attention. When the distance between motor vehicle and pedestrian arrives at traffic conflict threshold, the driver had to take immediate mitigation measures to avoid accidents. While the other way is: people, cars have been observed in time to each other, at this time the distance between the

two sides is longer than traffic conflict threshold, but the two participants do not have good communication. They are all thinking about whether the other party will allow himself to pass through. Then the situation occurs that the car will stop people leave when the pedestrian takes the same reaction. Both of these methods are common in road sections.

Therefore, if we want to improve the traffic environment and reduce the traffic conflicts, we must first communicate with pedestrians, for example, if the communication gestures designed in advance to indicate each other to understand each other's intentions, it is not easy to conflict, so that the traffic safety and efficiency would increase. In addition, pedestrian crossings will lead to a significant reduction in conflict, but also need to do a good job of traffic regulations and traffic safety knowledge of the publicity and popularization. In the section of the crosswalk, more eye-catching traffic signs are set to remind the driver in front of the existence of crosswalks and other measures also contribute to the reduction of conflicts and improve transport efficiency.

In addition, the road traffic environment design also has the following recommendations:

In the design process of the pedestrian crossing facilities, the arrival characteristics of pedestrians should be taken into full account to ensure the crosswalk's reachability. So as not to allow pedestrians to take shortcuts into the crosswalk, which could cause motor vehicle drivers unable to find pedestrian in time.

In order to reduce the probability of occurrence of conflict, it is necessary to ensure that the traffic participants can distinguish the potential hazards in time, so as to reduce the probability of conflict. The limited road resources in the various traffic participants (motor vehicles, non-motorized, and pedestrians) between the scientific and rational distribution.

In the current reality of the situation, the need for long-term strict management and active publicity and education, training road traffic in both community, mutual respect for each other right way to good habits, and establish a good road traffic participants communicate with each other mechanisms to ensure timely and efficient communication, understand each other's intentions.

Conclusion

In this paper, video was used to analyze qualitatively and quantitatively the pedestrian conflict of pedestrian crosswalks in typical urban road sections. After analyzing the conflict data, the corresponding relationship between the conflict distance and the conflict severity is established, which can be used as the basis for early warning of driver's dangerous driving behavior, so as to improve driver alertness and reduce the incidence of traffic accidents. And fundamentally improve the safety of the traffic system.

References

1. People's Republic of China Road Traffic Accident Statistics Compilation. 2003. Beijing: Ministry of Public Security Traffic Management Bureau.
2. Yoshida, T., H. Kuroda, and T. Nishigaito. 2004. Adaptive driver-assistance systems. *Hitachi Review* 19: 216–220.
3. Yuntong, Liu. 2004. *Road Traffic Safety Guide*. Beijing: People's Communications Press.
4. Zhang, Su. 1998. *Chinese Traffic Conflict Technology*. Chengdu: Southwest Jiaotong University Press.
5. Josephson, L., R. Weissleder, et al. 2000. Improvement of MRI probes to allow efficient detection of gene expression. *Bioconjugate Chemistry* 11(6): 941–946. ISSN: 1043-1802.
6. Khan, S.I., and W. Raksuntorn. 2001. Accuracy of numerical rectification of video images to analyze bicycle traffic scenes. *Transportation Research Record Journal of the Transportation Research Board* (3): 1773–1790. ISSN: 0361-1981.

Automotive Fire Simulation Based on Pyrosim

Hongguo Xu, Na Li, Hongfei Liu and Yihua Zhang

Abstract Automotive fire is one of the main forms of traffic incidents, and it is a theoretical and practical significance to analyze the automotive fire for reducing the number of fire. Through the application of fire process development algorithm, the material parameters of the materials are set up to give the model of automotive fire. The aim of using Pyrosim to simulate the fire model, obtain related data and analysis of fire smoke concentration and temperature in different space at the scene changes with time is to put forward the guiding suggestions for decreasing automotive fire, and for providing theoretical basis for optimization of the use and design of automotive.

Keywords Traffic safety · Automotive fire simulation · Pyrosim · Smoke · Temperature

Introduction

The development of science and technology in transportation has brought great changes to human life since the twenty-first century. But science and technology is a double-edged sword. The rapid development in transportation not only brings convenience but also produces many security risks, such as automobile fires, which

H. Xu · N. Li · H. Liu (✉) · Y. Zhang
School of Transportation, Jilin University, Changchun 130022, China
e-mail: hongfeiliu@jlu.edu.cn

H. Xu
e-mail: xuhg@jlu.edu.cn

N. Li
e-mail: lina19921988@163.com

Y. Zhang
e-mail: zhangyihua1989@126.com

are threatening the safety of people's lives and becoming one of the major disasters. According to the statistics released by Fire Department of United States [1], there are 152,300 car fires every year between 2006 and 2010, leading to the death of 209 people, 764 cases of injuries, and $536 million in direct property damage annually. In our country, there are 12,297 car fires every year since 2008, causing huge economic losses, posing a serious threat to human lives, and also putting a very negative impact on the development of the automotive industry [2].

The automotive fire model is based on Pyrosim, the aim of which is to analyze and predict the spread of smoke and temperature and the density of the toxic gases, hoping to provide guiding suggestion and theoretical basis for decreasing fire hazards.

Constructing Automotive Fire Model

Mesh

The mesh is made up of grids. The grid is the minimum and the basic unit of automotive fire model [3]. The calculation is based on the Poisson distribution with fast Fourier transform (FFT) method [4], so the grid's cell size should be in accordance with module 2l3m5n. The algorithm of the Pyrosim is to assign the obstruction calculating the individual cells. If the size you set of the obstruction is smaller than the grid size in automotive fire simulation, there will be some mistakes when the obstruction is diverted to the border of the grid.

The size of the grids decides the accuracy of results for vehicle fire simulation. The finer the grid is, the more accurate the results are. In the simulation based on the Pyrosim, the estimations are done in all grids within each time step, which means that the number of the grid plays an important role for the time of simulation. The finer the grid is, the longer the simulation time takes. There are three areas: the fire one, the buoyant plume district, and the smoke layer. The first is the most sensitive for the density of the grid, while the last is not as sensitive as others do. So to increase the number of grids in fire area and to reduce the number of grids in other areas reasonably [4, 5], can not only make the results accurate, but can also reduce the time taking. Therefore, the grid is divided into fire area and non-fire one [6].

Materials

To define the materials of the obstruction and the combustion characteristics of materials are the fundamentals of the simulation. Here, the materials are divided into seven sections, as listed in Table 1.

Table 1 Combustion characteristics of material

Material	Density (kg/m^3)	Specific heat (kJ/(kg·K))	Conductivity (W/(m·K))	Emissivity	Absorption coefficient	Application
Alloy steel	7850.0	0.46	45.8	0.95	5.0 E4	Metal parts
Painted steel	7930.0	0.50	45.8	0.92	5.0 E4	Bodywork
Foam	28.0	1.70	0.05	0.90	5.0 E4	Seat
Glass	2500	0.79	1.10	0.90	5.0 E4	Windows
Petrol	737.0	2.20	0.17	1.00	40.0	Fuel
Rubber	850.0	1.70	0.10	0.89	5.0 E4	Tires
Synthetic materials	1380.0	1.47	0.17	0.95	5.0 E4	Dashboard

Surface

Combustion characteristics of material surfaces are used to determine the obstruction surface, while the set of the surface is based on the material. In other words, the surface of obstruction is determined by the material.

Fire

Due to the principle of Pyrosim is to break each obstruction estimated on the border to the grid, and fire belongs to vent group, which can only be defined by two-dimensional one. Therefore, the fire must be set in the grid on the border, or there will be an error during the run-time for the vent cannot be break down to the border. So, it is the key to have an accurate judgement about the border to the grid for setting fire in the simulation.

The surface of fire is the fundamental to set fire. There is something which is really different from defining other surfaces, for it is no need to depend on the material, but to define the combustion of the fire directly. The heat release rate (HRR) of fire is 1000.0 kW/m^2.

Obstruction

Obstruction is the most important appearance of the composition. Relatively to the FDS, the model in Pyrosim has been improved much, but compared to other professional graphics software, there are still many shortcomings. For example, rectangular obstruction could be given directly, while additional obstructions of which have non-rectangular borders could not be drawn directly [7]. But tires are based on cylinder, so it is necessary to take another approach to model tires. Here are three ways to solve the issue [8].

Tiny Elements Construction

Circle can be seen as constructed by many rectangles. The size can be obtained by using a number of parallel lines to cut the circle, as shown in Fig. 1, and a cylinder can be given in the same way, too. But the boundaries are very rough, while it can be smoothed by refining the cubes and increasing the number of them, as shown in Fig. 2, but it will lead to a lot of work. So, it is not an ideal method.

Copy and Rotation

A cube can be modeled first, and then it can be copied, during rotating it around a specific point inside or outside of the cube to give a cylinder or a hollow cylinder, as shown in Figs. 3, 4, and 5. The cylinder not only has smooth boundary but also reduces the workload. Last but not least, though the cylinder is made of many cubes, it has no effect on combustion characteristics.

Fig. 1 Circle model
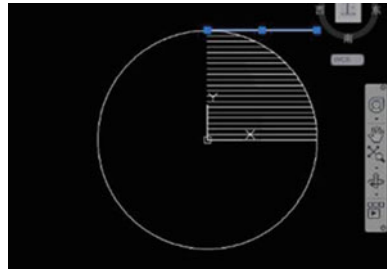
Fig. 2 Cylinder model
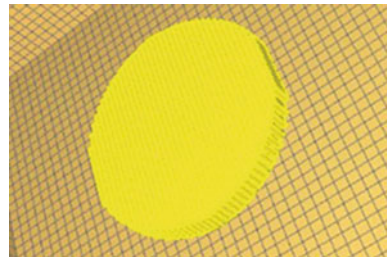
Fig. 3 Constructed cube
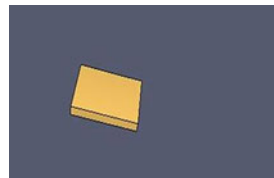
Fig. 4 Copy and rotation
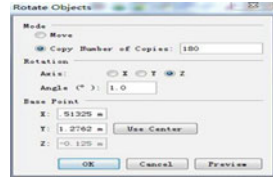
Fig. 5 Copy and rotation
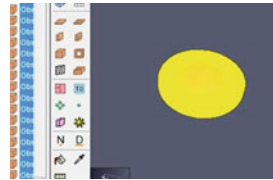
Fig. 6 Tire CAD sketch
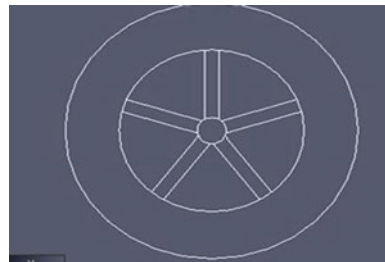
Fig. 7 Completed tire
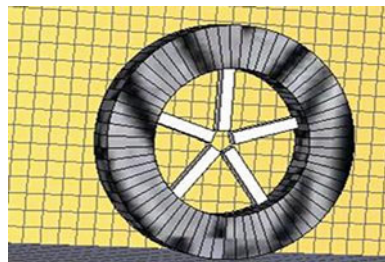
Import the File in dwf

Though the above two methods exist, tire is not a simple cylinder. If a tire is constructed by the two methods directly, it will greatly increase the amount of computation. Therefore, to input the file in dwf seems to be the best choice. First of all, it is necessary to draw the exterior contour of the obstruction in CAD, and then save the file in dwf format. Then, the file is an input to Pyrosim, as shown in Fig. 6. Finally, the lines are converted to walls, and the materials are set to get the following models, as shown in Fig. 7.

Device	Location		
	X(m)	Y(m)	Z(m)
Thermocouple 01	0.5	0.9	0.0
Thermocouple 02	0.5	0.9	0.2
Thermocouple 03	0.5	0.9	0.4
Thermocouple 04	0.5	0.9	0.6
Thermocouple 05	0.5	0.9	0.8
Thermocouple 06	0.5	0.9	1.0
Smoke detector	2.0	0.8	1.1

Table 2 Position of the thermocouple when the burner surface is located in the engine area

Device

Here are many devices in Pyrosim, such as thermocouple, flow measuring device, HRR device, and beam detector device. Here, the first two are chosen, thermocouple and flow measuring device. A group of thermocouples are set near the fire, and they are kept at a certain distance in height [9] to acquire the temperature distribution in different height at different times. The set of smoke detector is based on the safety of the occupants. It is made above the seat of the driver. The details are shown in Table 2.

Slice of Temperature

The slice of temperature in Pyrosim is a specific plane of the grid. It can show the temperature of all the points of the plane in the simulation process at different times. So the dynamic changes can be seen to go deep into the study. That has great importance. Here are four slices in the simulation, including $y = 0.9, 0.5, x = 0.8, 1.0, 1.7, 4.1$.

Parameters

Smoke view is essential to realize the simulation of Pyrosim. The computer will run according to the default time step with 0.01 s, which calls for powerful computer configuration, and a lot of time. So, increasing the time step is very necessary. A total of 10.00 s are chosen [10] as the time step, which has minimal effect on the result, and it ends the simulation at 2000.00 s.

Model

The model is shown in Figs. 8 and 9.

Fig. 8 Completed model when the burner surface is located in the engine area
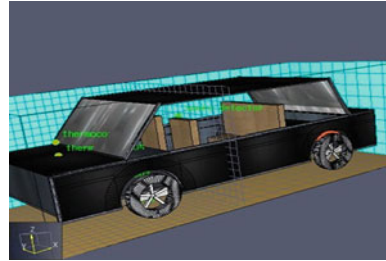
Fig. 9 Completed model with slices
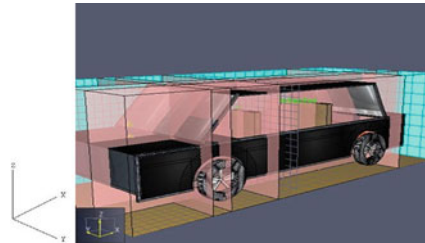
Fig. 10 Results of smoke detector
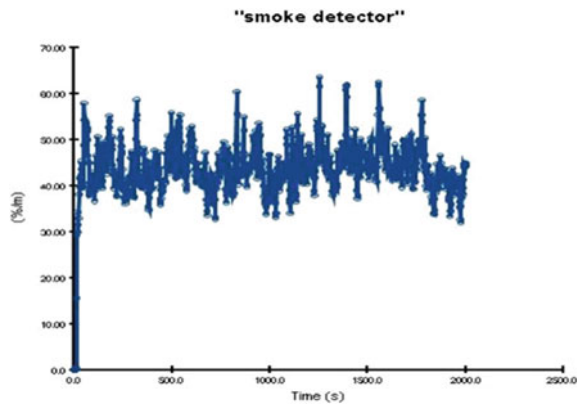
Analysis of Simulation Results

The Result of the Smoke Detector

According to the statistical analysis of the cause of death of people in automotive fire, the most damage is caused by smoke. Nearly 78.9% of the deaths are due to the smoke [11]. So it is very necessary to analyze the smoke concentration. The result of smoke detector is shown in Fig. 10.

It can be seen from the figure that the smoke concentration is really high when the fire is on the engine area. There will be a great threat to the safety of passengers if exposed to such environment for a certain time. So something must be done to prevent it.

The Results of the Thermocouples

The results of the thermocouples are shown in Figs. 11, 12, 13, 14, 15, and 16.

Thermocouple 01, Thermocouple 02, and Thermocouple 06 are external to the automotive body; because of the existence of the body and the chassis, the heat transfer is hampered. Although there are many large fluctuations in Figs. 11, 12, and 16, it is just fluctuating around 20 °C. It is very close to the environment temperature, which means that in 2000 s, the outside of the body almost has no influence.

Thermocouple 03 and Thermocouple 04 are under the burner, as shown in Figs. 13 and 14. Although the temperature is not very high, it has a gradually rising trend. Thermocouple 05 is above the burner and closer to it, so the temperature is higher. It should be noted that the HRR is rising very fast to reach the criteria set in 1.0 s by default. So the temperature of Thermocouple 05 is rising rapidly at the

Fig. 11 Results of Thermocouple 01
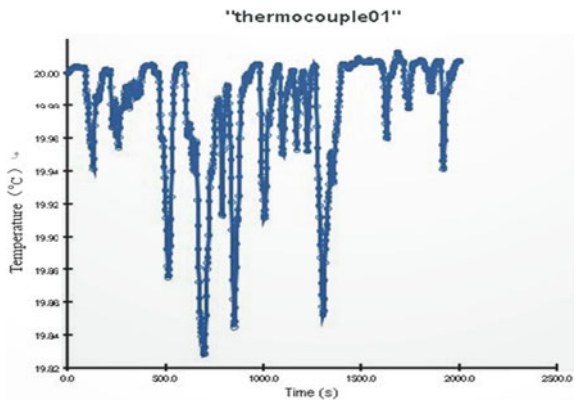
Fig. 12 Results of Thermocouple 02
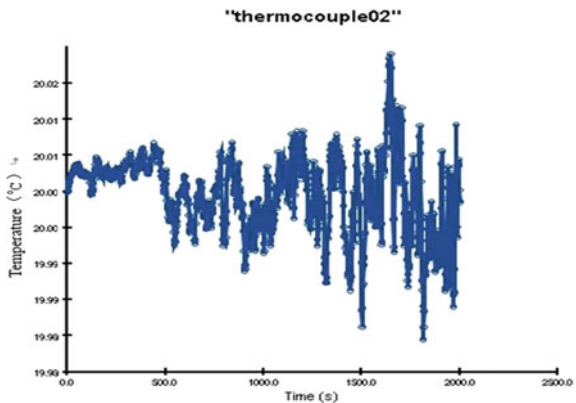
Fig. 13 Results of Thermocouple 03
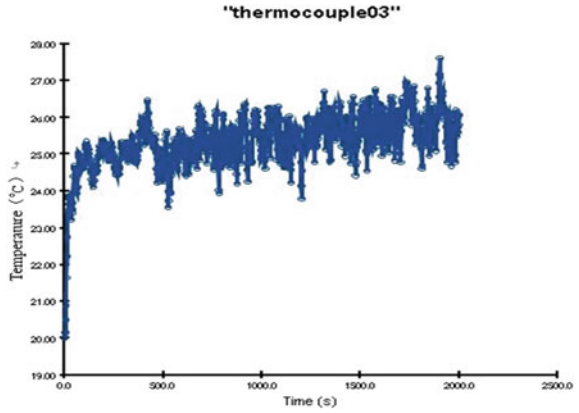
Fig. 14 Results of Thermocouple 04
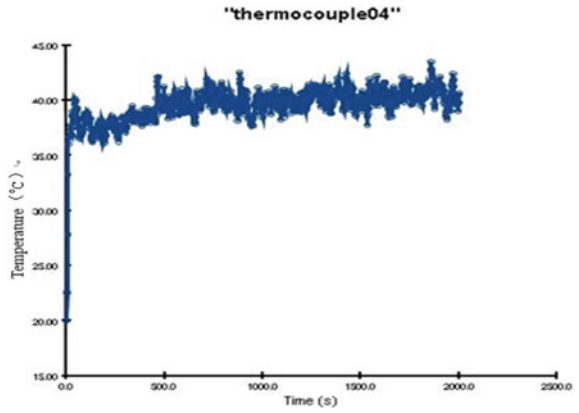
Fig. 15 Results of Thermocouple 05
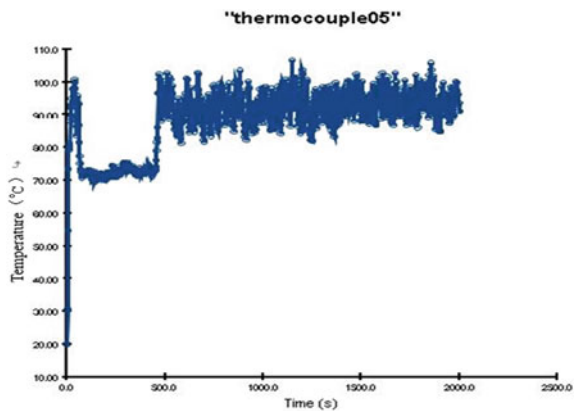
Fig. 16 Results of
Thermocouple 06
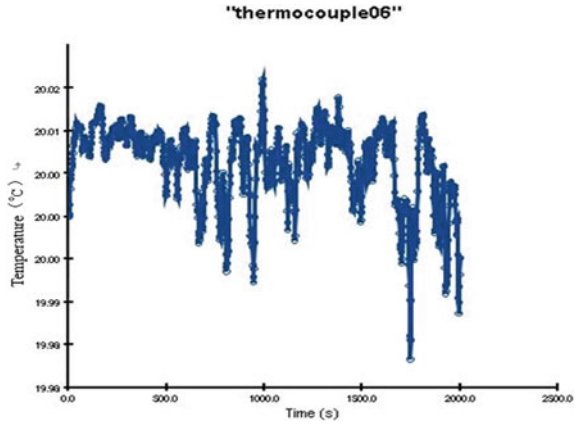
beginning of the simulation. Then, it declines nearly 30 °C in a short time and
keeps for a period of time, because in the initial stage of combustion, the burner will
absorb the heat around the fire to sustain burning. At the end of this period, the
temperature starts to rise again.

According to the above analysis, the distance has a great effect on the temperature. Closer away from sources of ignition, the higher the temperature is. In
addition, alloy steel body has great resistance for thermal transfer and heat release.

The Results of the Temperature Slice

When the fire happens in the engine area, here are six slices around 1100 s, as
shown in Figs. 17, 18, 19, 20, 21 and 22. The blank space in the figures is where the
obstruction is, so the temperature could not be obtained.

Fig. 17 Thermogram of
temperature slice as
$x = 0.8$ m
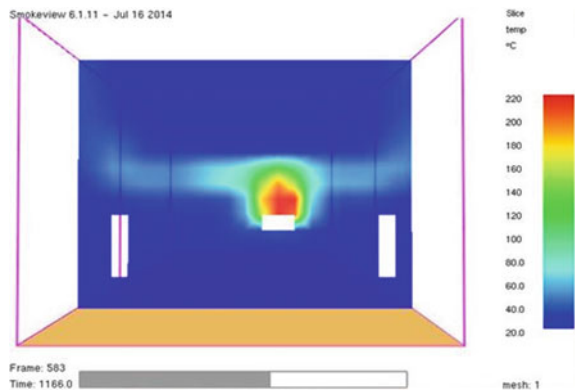
Fig. 18 Thermogram of temperature slice as $x = 1.0$ m
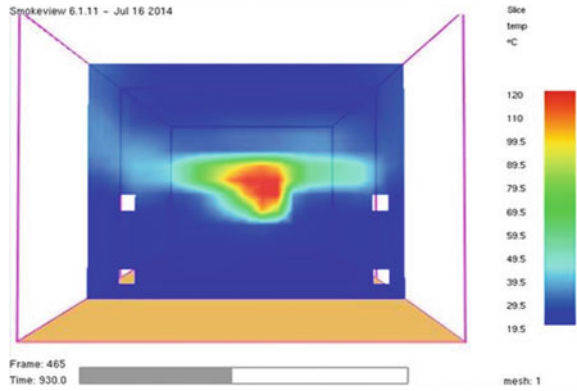
Fig. 19 Thermogram of temperature slice as $x = 1.7$ m
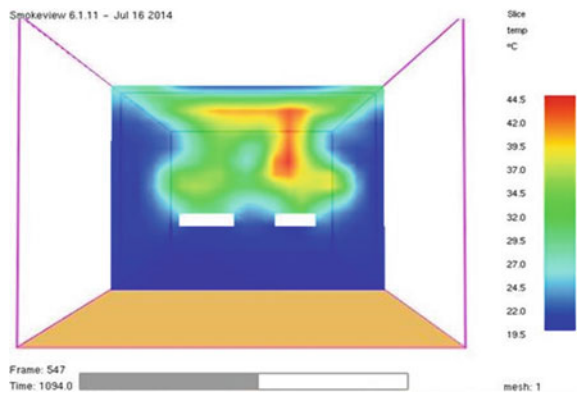
Fig. 20 Thermogram of temperature slice as $x = 4.1$ m
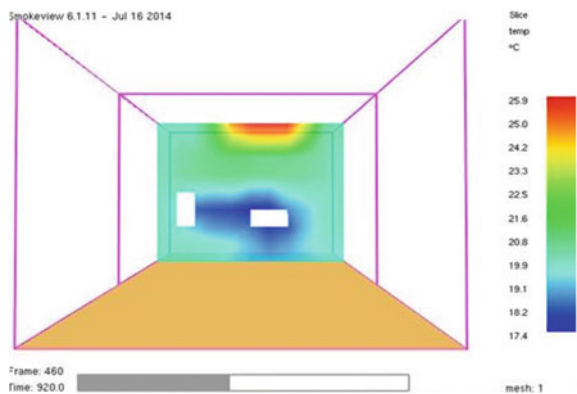
Fig. 21 Thermogram of
temperature slice as
$y = 0.5$ m
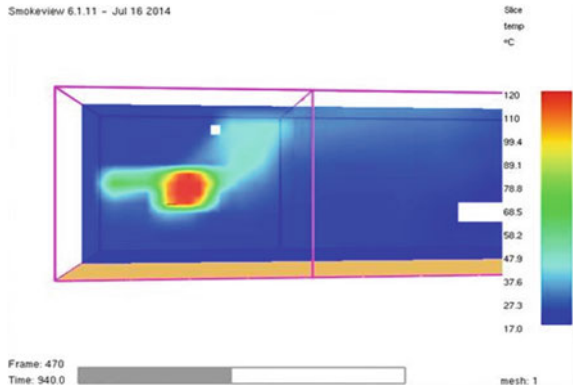
Fig. 22 Thermogram of
temperature slice as
$y = 0.9$ m
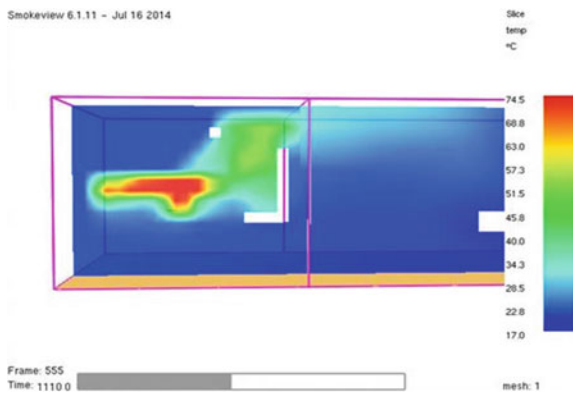
Setting the slices for they are special in the space, where $x = 0.8$ m is the right boundary of the burning area in the engine area; where $x = 1.7$ m is the middle plane of the drivers' seat on the direction of the X-axis; where $y = 0.5$ m is the middle plane of the drivers' seat on the direction of the Y-axis; where $y = 0.9$ m is the middle plane of the car on the direction of the Y-axis.

As shown in Fig. 17, $x = 0.8$ m is the right boundary of the burning area in the engine; at the beginning of the 500 s, flame keeps smooth. After 500 s, flame begins to fluctuate dramatically. Reasons for such a change may be due to the synthetic material below the burner is burning out, and petrol is going to burn, so the HRR of the fire is increasing rapidly, so the temperature does. Besides, the highest temperature is equal to 220 °C in the slice, which can be harmful to the occupants.

As shown in Fig. 19, $x = 1.7$ m is the middle plane of the drivers' seat on the direction of the X-axis, so it has great importance on the analysis. The highest temperature in the slice is 44.5 °C, though it would not cause harm to the occupants, and it is also a burden to the safety of the human. But as time goes by, the temperature in the slice has a rising trend. If occupants are exposed to such

environment, the possibility of harm to the human body will increase rapidly [12]. While the temperature in the slice where $x = 1.0$ m, $x = 4.1$ m is not very high, which means when the burner is in the engine area, it has little effect on the dashboard and luggage compartment [13].

$y = 0.5$ m is the middle plane of the drivers' seat on the direction of the Y-axis, as shown in Fig. 21. The highest temperature is equal to 74.5 °C, as shown in Fig. 19. It is high enough to do harm to the human body. These facts show that the farther the burner is, the lower the temperature will be. And the high temperature point offset upward from the center of the area. This also proves that hot air has a tendency to rise.

Conclusion

The following conclusions were drawn from the above work:

1. The distance between the fire and the passenger is the most important factor related to safety. The closer the passenger and fire is, the faster the temperature of the crew area rises, and the higher concentrations of smoke rises, which means the harm to humans will occur earlier.
2. The air has the trend to be upward when heated, so as the smoke. Therefore, passenger should properly reduce the body center of gravity and keep calm to find reasonable ways to survive.
3. Materials play a very important role in fire simulation. As different materials have many differences in combustion characteristics, they always have great impact on the spread of fire and the heat transfer. So there should be a reasonable choice of body and the internal structure of materials, to the greatest extent possible to reduce the losses on the economy and security of the people caused by automotive fire.

References

1. Ahrens, M. 2012. Automobile fires in the U.S: 2006–2010 estimates. In *Fire in Vehicles (FIVE) Conference in Chicago*, Illinois, Sep 27, 2012.
2. Zhang, Y. F., H. Huang, and Cao Liying. 2014. The analysis and research of automobile fire in China. In *Fire-proceedings of the annual meeting of the association of science and technology in China*.
3. Qun, Y., L. Chang, and Z. Liang. 2008. Analysis on fire simulation and application based on FDS. *Journal of water resources and construction* (4): 124–126.
4. Mcgrattan, K.B. 2004. *Fire Dynamics Simulator (Version 4) Technical Reference Guide*. NIST Special Publication 4(Jan/Feb), 206–207.
5. Zhu, S., R. Huo, and L.H. Hu., et.al. 2008. The Impact of extension of the calculation mesh and opening of regional on FDS simulations. *Journal of safety and environment* (4): 131–135.
6. Hongfei, Wei. 2015. A simulation of the fire in a students' dining room. *Safety* 36 (3): 9–11.

 7. Xue, Wei, and Guangjun Zhang. 2006. FDS fire simulation and application. *Jilin Forestry Science and technology* 35 (6): 18–20.
 8. Shang, C., K. Wang, and H. Huang, et al. 2013. The construction of trees in fire based on pyrosim. *Fire science and technology*, 32 (9): 1030–1033.
 9. Liang, J.H., X. Li, and J.H. Lin, et al. 2014. Analyses of EMU fire based on pyrosim. In Rail Passenger Safety in Fire Safety and Fire-Retardant Technology Symposium Proceeding.
10. Chen, W., H.H. Cui, and L. Qin, et.al. 2011. Research of fire simulation based on FDS. *Computer simulation* 28 (12): 227–231.
11. Zhang, Z., and J. Zhang. 2015. On the laboratory building fire and safe evacuation based on computer simulation. *Value engineering* (7): 308–309.
12. Liu, Y. 2012. Study of Numerical Simulation and Safe Evacuation for Building Fire Based on FDS. China: Liaoning technical university.
13. Li, Y. 2015. Based on number of FDS simulation gao tiezhong luggage influence on fire hazard. *Fire technique and products information* (4).
14. Liu, J., M. Liu, and H. Zhi, et al. 2006. FDS fire simulation analysis on basic theory and application techniques. *Safety*, 27 (1): 6–9.
15. Gang, Shao, Yang Peizhong, and Jin Xianlong. 2005. Automatic modeling for tunnel with non-rectangular wall in FDS. *Computer engineering and applications* 41 (36): 213–216.

Passenger Flow Distribution Model Under the Interruption of Urban Rail Transit Network

Xue Han, Di Wang, Yingshun Liu and Tangyi Guo

Abstract Urban rail transit has gradually become an important part of urban public transport system due to its advantages of large capacity, punctuality, comfortableness, and environmental protection. As the growing popularization of urban rail transit network, a large number of passengers will be stranded once the rail transit line is interrupted. In the context of interruption of urban rail transit network, a NL model for passenger travel selection has been established under the circumstance of network interruption for passenger flow redistribution by determining effective passenger flow share rate of each route according to the passenger flow capacity of such route. It has been demonstrated that passenger flow share rates of metro, bus, and taxi are 50.7, 30.46, and 17.81%, respectively, by adoption of the proposed model.

Keywords Urban rail transit · Line interruption · Passenger flow distribution · NL model

Introduction

In urban rail transit network, various emergency occurrences may interrupt transit lines, whereby leading to passenger stranding and congestion. Therefore, under the circumstance of urban rail transit network interruption, it is the fundamental aspect for passengers evacuation and traffic congestion relief, and also an important basis for reasonable transport resource distribution and urban rail transit service improvement, to study passengers' travel route selection, analyze major factors that influence their selections and predict the passenger share rates of various transportation routes.

X. Han (✉) · D. Wang · Y. Liu · T. Guo
Hohai University Wentian College, Manshan 243031, Anhui, China
e-mail: hanxue_echo@163.com

Much of the studies about passenger flow distribution in rail transit have adopted the theory of urban road traffic passenger flow distribution [1]. Guo and Shen [2] proposed to divide transport distribution models by joint method of cooperative and competitive ODs, and also conducted deep analysis for transport distribution division by layered strategic transport method and joint method. Wu and Liu [3] used Wardrop Equilibrium Principle and Frank–Wolfe algorithm for the solution of that model, having achieved the distribution of passenger flow in rail transit lines. However, the assumption of Equilibrium Principle is that all the used routes have the minimum and equal impedance, and the impedance of unused routes is larger than the minimum, which has limited the application of such principle. Later, Zeng et al. [4] have improved Wardrop Equilibrium Model by considering factors influencing passenger flow distribution, and achieved the passenger flow distribution by converting passengers' psychological and cost factors into passenger satisfaction rate. In addition, Si et al. [5] built a passenger flow distribution model with equilibrium random passengers who make transfers, by taking account of features of urban rail transit network. Xu et al. [6] built a probability model in compliance with normal distribution for passenger flow distribution by combining characteristics of urban rail transit and impedance function of each route. Zhang [7] has made improvement for the optimal combination model in terms of strategic improvement and combination load against the weaknesses of EMME model, making it be more flexible to practical cases. Song et al. [8] added saturation entropy to broad travel cost to form a more reasonable impedance function, and used chaos optimization algorithm for solution, whereby distributing passenger flow in comprehensive transit network. It is difficult to accurately search passenger's route selection information by using AFC system because of the seamless transfer condition in urban rail transit network; therefore, Lu and Liu [9] introduced attraction coefficient algorithm to simulate passenger's route selection probability. Lin et al. [10] proposed an improved method of urban rail transit passenger flow distribution by establishing a broad cost–benefit analysis model and combining passenger route selection probability based on random utility theory. Ii has been verified that this model can better simulate passenger route selection in urban rail transit network.

Factors Influencing Stranded Passengers' Route Selection in Urban Rail Transit

Compared with other transport methods, urban rail transit has the advantages of fixed lines, not easily being influenced by external factors and good punctuality. Therefore, factors influencing passengers' route selection are somewhat different. In addition, passengers will be emotionally influenced when metro line is interrupted, being more irritable.

(1) Travel time

Travel time in rail transit line could be defined as the total time from passenger enters origin station to passenger leaves the destination station, including vehicle travel time, transfer time, transfer waiting time, and vehicle stopping time. Generally, passengers care more about travel time and cost. However, when traffic line is interrupted, passengers are more sensitive to time issue, wishing to reduce waiting time, to leave the affected station, and to reach destination as soon as possible. Hence, if various routes are available, passengers will choose the quickest one.

(2) Transfer time

Transfer time is divided into transfer walking time and transfer waiting time. The former refers to the time needed for one passenger walking from station of one rail line to the station of another rail line. The latter refers to the time needed for one passenger to step into the to-be-transferred vehicle after reaching corresponding station. In addition, transfer walking time is affected by external factors such as distance between transferred stations, staircase steps, and the availability of escalator, and internal factors such as passenger's own physical condition, walking speed, and station familiarity.

(3) Transfer times

Transfer times are the total number needed by one passenger to finish his/her travel. Survey on Passenger Travel through Shanghai Metro has shown that most passengers are tolerant of three transfers, the travel route with less than three transfers are most preferred. Because transfer times can directly reflect the convenience of specific route, passengers generally take transfer times as an important consideration in route selection.

(4) Travel cost

Travel cost is another important factor for passenger selecting travel route. However, under the condition of rail line interruption, passengers desire to reach destination as soon as possible, thus being less sensitive to travel cost. Despite of this, travel cost plays an important role in passenger travel route selection if various charge systems are available. There are three charge systems for China metro, including flat fare, charge by stations, and charge by distance. Because of small ticket price fluctuation in flat fare and charge by station, influence of price on passenger route selection is less significant. While passengers generally choose route with the shortest travel distance under charge by distance system in order to save travel cost.

(5) Network familiarity

Influence of network familiarity on passenger route selection is mainly demonstrated as whether passenger can choose more convenient traffic route according to his/her own understanding of traffic network and travel experience.

Generally, commuters select the route consuming less time and of more convenience because of their relatively fixed travel routes and rich experience. Nevertheless, when passengers are not familiar with network, they often choose their familiar route which is not always the shortest one.

NL Model of Passenger Flow Redistribution

Random Utility Theory

Disaggregate model is developed on the base of "maximization of utility" principle [11] in economics, namely consumers are rational persons who will choose the commodity or commodity combination that can most meet their needs. Therefore, the following assumptions should be met during the establishment of passenger utility function.

(1) Passenger is the subject to make travel selection and will make such travel decision independently.
(2) Passenger will make travel plan that meets his/her travel needs in accordance with maximization of utility principle.

Therefore, utility function U_{in} for plan i chosen by passenger n could be given as [11]

$$U_{in} = V_{in} + \varepsilon_{in} \tag{1}$$

in which the fixed term of this utility function for plan i chosen by passenger n is represented as V_{in}, while the probability terms is represented as ε_{in}.

Based on maximization of utility theory, the probability P_{in} for passenger n to choose plan i could be given by Eq. (2)

$$\begin{aligned} P_{in} &= \mathrm{Prob}(U_{in} > U_{jn}; \quad i \neq j, \ j \in A_n) \\ &= \mathrm{Prob}(V_{in} + \varepsilon_{in} > V_{jn} + \varepsilon_{jn}; \quad i \neq j, \ j \in A_n) \end{aligned} \tag{2}$$

in which V_{jn} is the fixed term of utility function that means passenger n will choose plan j, while ε_{jn} is the probability term of such function. $0 \leq P_{in} \leq 1$; $\sum_{i \in A_n} P_{in} = 1$, A_n is the set of plans for passenger n.

NL Model Establishment

Assuming that there is a double-layered selection tree structure, the upper selection layer set 2 has T plans, while the lower selection layer set 1 has M plans, as shown in Fig. 1 [12]
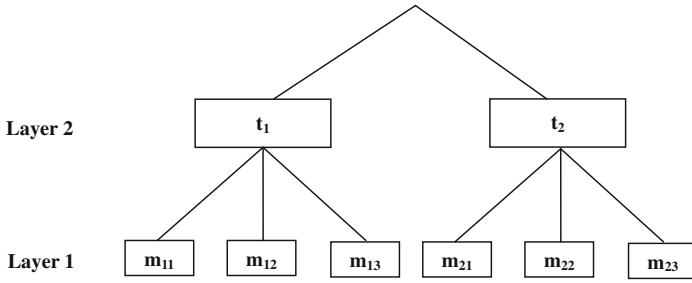
Fig. 1 Diagram for NL model selection tree

In which $T = (t_1, t_2, \ldots, t_i)$ $M = (m_1, m_2, \ldots, m_j)$. The selection set pf plans can be expressed as

$$
\begin{aligned}
A &= T \times M \\
&= \left\{ (t_1 \times m_1), \ldots, (t_1 \times m_j), (t_2 \times m_1), \ldots, (t_2 \times m_1), (t_i \times m_1) \ldots (t_i \times m_j) \right\}
\end{aligned}
\tag{3}
$$

Total utility of plan A can be represented as

$$
U_{tm} = V_t + V_m + V_{tm} + \varepsilon_t + \varepsilon_m + \varepsilon_{tm}
\tag{4}
$$

Assuming ε_{tm} and ε_t follows double exponential distribution with variation coefficients of λ_m and λ_t, and then, according to maximization of utility theory, the marginal probability for passenger n to choose plan t is [12]

$$
\begin{aligned}
P_n(t) &= \text{Prob}(\max_{m \in M} U_{tm} > \max_{m \in M} U_{t'm}, \quad \forall t \in T, \ t' \neq t) \\
&= \text{Prob}\left\{ V_t + \varepsilon_t + \max_{t \in T}(V_t + V_{tm} + \varepsilon_{tm}) \geq V_i + \varepsilon_i + \max_{t \in T}(V_{t'} + V_{t'm} + \varepsilon_{t'm}) \right\} \\
&= \frac{\exp\{(\beta_t X_t + V_t)\lambda_t\}}{\sum_{t \in T} \exp\{(\beta_t X_{t'} + V_{t'})\lambda_t\}}
\end{aligned}
\tag{5}
$$

Making $V_i = \ln \sum_{m \in M} \exp(\beta_t X_t + \beta_{tm} X_{tm})$, it could be inferred according to conditional probability formula $P_n(t, m) = P_n(m|t) \cdot P_n(t)$ that

$$
P_n(t, m) = \frac{\exp(\beta_m X_m + \beta_{tm} X_{tm})}{\sum_{m \in M} \exp(\beta_m X_{m'} + \beta_{tm} X_{tm'})} \cdot \frac{\exp(\beta_t X_t + \lambda_t V_t)}{\sum_{t \in T} \exp(\beta_t X_{t'} + \lambda_t V_{t'})}
\tag{6}
$$

in which V_t is the fixed term of utility function for passenger to choose plan t; V_m is the fixed term of utility function for passenger to choose plan m; V_{mt} is the fixed term of utility function for passengers to choose plan m and plan t simultaneously; ε_t is the probability term of utility function for passenger to choose plan t; ε_m is the probability term of utility function for passenger to choose plan m; ε_{mt} is the probability term of utility function for passenger simultaneously to choose plan m and plan t. β_t, β_m, and β_{tm} are solve for parameters, λ_t represents inclusive coefficient; X_t, X_m, and X_{tm} represent the sets of influencing factors when passenger choose plan T, plan M, and both plan T and plan M, respectively.

Applied Case

SP Survey

SP survey has the advantages of small survey scope, high feasibility, small errors, and concentrated selection plans. Therefore, in order to obtain travel selection data of Shanghai citizens, both paper and online questionnaires have been used in this paper. A total of 500 questionnaires (200 online questionnaires) were released, of which 226 effective ones were returned with 904 experimental values. The starting points designed in the questionnaire were Shanghai Railway Station and Oriental Sports Center. The metro line interruption occurred in People's Square Station, which included three transferred railway lines, two transferred bus lines, and one taxi line. The information about specific time, transferred times, and cost of the six situations is shown in Table 1. A total of 100 returned effective questionnaires were randomly sampled for further analysis.

Table 1 Travel selection situation design under the circumstance of Shanghai metro line interruption

	Selection plan	Total travel time (min)	Transfer time (min)	Travel time (min)	Cost (RMB Yuan)	Transfer times
Metro	1	51	4	47	5	1
	2	50	4	46	6	1
	3	44	5	39	5	2
Bus	4	83	3	80	4	1
	5	94	4	90	6	2
Taxi	6	25	0	25	40	0

Determination of Characteristic Variables and Utility Functions

When urban track transit line is interrupted, passengers will consider to take other metro lines, buses, or taxi to reach their destinations. According to situation design and practical conditions of Shanghai transport, the proposed NL model structure has two layers as shown in Fig. 2.

According to the above analysis for factors influencing route selection of stranded passenger after track transit interruption, the characteristic variables of NL model are travel time (TT), transfer waiting time (WT), transfer times (HC), and travel cost (C), as shown in Table 2.

Utility function of each selection branch established according to Table 2 is shown as follows:

$$V_{rmn} \equiv V_{(r|m)_n} + V_{mn} \tag{7}$$

$$V_{(r|m)} = \beta' X_{(r|m)_n} = \sum_{k=1}^{K_1} \beta_k X_{(r|m)_n} \tag{8}$$

$$V_{(mn)} = \beta' X_{mn} = \sum_{k=1}^{k_2} \theta_k X_{mn} \tag{9}$$

in which $\beta' = [\beta_1, \ldots, \beta_K, \ldots, \beta_{K_1}]$ represents the unknown parameter vector of K_1 order when layer 1 is selected.

$X_{(r|m)_n} = [X_{(r|m)_{n1}}, \ldots, X_{(r|m)_{nk}}, \ldots, X_{(r|m)_{nK_1}}]'$ represents characteristic variable of K_1 order.
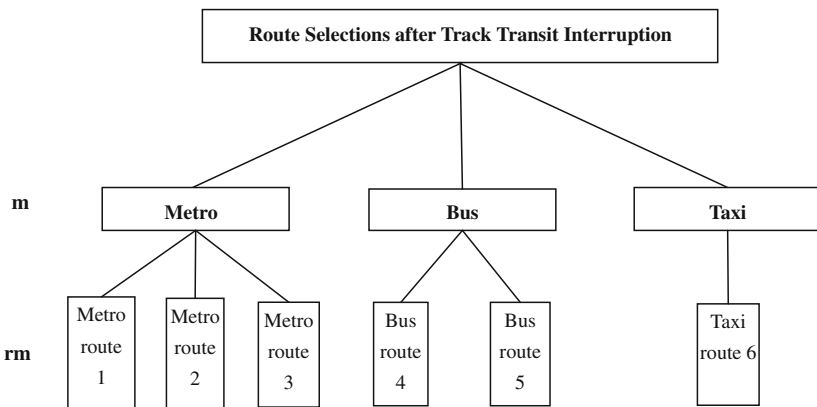
Fig. 2 Structure of NL model for route selections after Shanghai track transit interruption

Table 2 Characteristic variables of NL model

Selection plan		Characteristic variables			
		Travel time	Waiting time	Transfer times	Travel cost
Metro	1	$TT_{(1|m)}$	$WT_{(1|m)}$	$HC_{(1|m)}$	$C_{(1|m)}$
	2	$TT_{(2|m)}$	$WT_{(2|m)}$	$HC_{(2|m)}$	$C_{(2|m)}$
	3	$TT_{(3|m)}$	$WT_{(3|m)}$	$HC_{(3|m)}$	$C_{(3|m)}$
Bus	4	$TT_{(4|b)}$	$WT_{(4|b)}$	$HC_{(4|b)}$	$C_{(4|b)}$
	5	$TT_{(5|b)}$	$WT_{(5|b)}$	$HC_{(5|b)}$	$C_{(5|b)}$
Taxi	6	$TT_{(6|t)}$	$WT_{(6|t)}$	$HC_{(6|t)}$	$C_{(6|t)}$

$\theta' = [\theta_1, \ldots, \theta_k, \ldots, \theta_{k_2}]$ represents the unknown parameter vector of K_2 order when layer 2 is selected.

$X_{mn} = [X_{mn1}, \ldots, X_{mnk}, \ldots X_{mnk_2}]'$ represents characteristic variable vector of K_2 order.

Taking the above equation into conditional probability equation:

$$P_n(rm) = P_n(r|m) \cdot P_n(m)$$
$$= \frac{e^{\lambda_1 \beta' X_{(r|m)n}}}{\sum_{r'=1}^{r_{mn}} e^{\lambda_1 \beta' X_{(r|m)n}}} \cdot \frac{\exp[\lambda_2 \theta' X_m + (\lambda_2/\lambda_1) \ln \sum_{r'=1}^{r_{mn}} \exp(\lambda_1 \beta' X_{(r'|m)_n})]}{\sum_{m'=1}^{m_n} \exp[\lambda_2 \theta' X_m + (\lambda_2/\lambda_1) \ln \sum_{r'=1}^{r_{mn}} \exp(\lambda_1 \beta' X_{(r'|m)_n})]}$$

(10)

Parameter Calibration and Test

MATLAB programming was used to calibrate and test parameters for above data. Calibration value and test value of each parameter in NL model are listed in Table 3.

It can be seen that the signs of all parameters in travel attribute row are negative, indicating travel time, transfer waiting time, transfer times, and travel cost will have negative utility on route selection behavior, which are in compliance with practical conditions. Absolute value of T-test result for each parameter is larger than 1.96. Therefore, it can be estimated in a 90% probability that each parameter plays a significant role in travel route under the interruption of track transit line. ρ^2 is larger than 0.2, indicating that this model is of high accuracy.

Taking the above calibration results into Eq. (7), utility function of each route selection can be obtained:

$$V_{(1|m)} = -0.712 - 0.0225TT - 0.29WT - 0.382HC - 0.0642C \qquad (11)$$

Table 3 Calibrating results of parameters in NL model

Variable type	Variable name		Parameter	Estimated value	Standard error	T-test value
Constant term	Metro constant	Route 1	ASC1	−0.712	0.834	−4.08
		Route 2	ASC2	0.224	0.128	5.31
		Route 3	ASC3	0.332	0.351	2.03
	Bus constant	Route 4	ASC4	0.365	0.810	4.79
		Route 5	ASC5	−0.217	0.124	−3.46
	Taxi constant	Route 6	ASC6	0.00842	0.183	−5.69
Travel attribute	Travel time		β_1	−0.0225	0.011	−2.32
	Waiting time		β_2	−0.29	0.565	−3.52
	Transfer times		β_3	−0.382	0.129	−2.34
	Travel cost		β_4	−0.0642	0.038	−5.08
Other test values	$L(0)$			343.957		
	$L(\theta)$			179.143		
	$L(0) - L(\theta)$			164.814		
	ρ^2			0.255		

Table 4 Share rate prediction for each travel route when urban track transit line is interrupted

Selection plan		Total travel time (min)	Transfer times	Cost (RMB Yuan)	Selection probability (%)	Share rate of travel method (%)
Metro	1	51	1	5	2.18	50.70
	2	50	1	6	5.20	
	3	44	2	5	43.32	
Bus	4	83	1	4	20.04	30.46
	5	94	2	6	10.42	
Taxi	6	25	0	40	17.81	17.81
Others	–	–	–	–	1.03	1.03

$$V_{(2|m)} = 0.244 - 0.0225\text{TT} - 0.29\text{WT} - 0.382\text{HC} - 0.0642C \qquad (12)$$

$$V_{(3|m)} = 0.332 - 0.0225\text{TT} - 0.29\text{WT} - 0.382\text{HC} - 0.0642C \qquad (13)$$

$$V_{(4|b)} = 0.365 - 0.0225\text{TT} - 0.29\text{WT} - 0.382\text{HC} - 0.0642C \qquad (14)$$

$$V_{(5|b)} = -0.217 - 0.0225\text{TT} - 0.29\text{WT} - 0.382\text{HC} - 0.0642C \qquad (15)$$

$$V_{(6|t)} = 0.00842 - 0.0225\text{TT} - 0.29\text{WT} - 0.382\text{HC} - 0.0642C \qquad (16)$$

A total of 100 questionnaires were randomly selected from the 226 effective returned ones in SP survey as predication samples. By taking investigation data into the above utility functions and making use of Eq. (15), share rate of each travel

method when urban track transit line is interrupted can be predicated, as shown in Table 4.

It can be obtained from Table 4 that when urban track transit line is interrupted, over 50% passengers will choose to transfer to other metro lines; 30.46% passengers to bus lines, 17.81% to taxi, and 1.03% to other routes or do not travel at all.

Conclusions

Through analyzing behavioral features of stranded passengers when urban track transit line is interrupted and in combination with investigation data, it has been judged that factors influencing route selections of stranded passengers are travel time, transfer time, transfer times, travel cost, and passengers' familiarity with track transit network. Based on the analysis of data from survey on travel selection intentions of Shanghai citizens and by establishing a disaggregated NL model of track transit line interruption, it has been demonstrated that travel time, transfer waiting time, transfer times, and travel cost are the major factors influencing travel method selections of passengers. Moreover, test analysis has been done to determine the accuracy of such proposed mode, showing that it is feasible to predict travel route selections of passengers under the circumstance of track transit line interruption using this NL model. Furthermore, it has been predicted that the share rates of metro, bus, and taxi are 50.7, 30.46, and 17.81% respectively by adoption of the NL model.

References

1. Wardrop, J. G. 1952. Some theoretical aspects of road. *Traffic Research* 02.
2. Guo, Xiucheng, and Lu, Shen. 2000. Study of URT's joint modal split assignment model on cooperative and competitive OD matrix. *China Journal of Highway and Transport* 13 (4): 91–94.
3. Wu, Xiangyun, and Canqi, Liu. 2004. Traffic equilibrium assignment model specially for urban railway network. *Journal of Tongji University (Natural Science)* 32 (9): 1158–1162.
4. Zeng, Mingkai, Jian, Huang, and Qiyuan, Peng. 2006. Research on assignment of passenger train plan for dedicated passenger traffic lines. *Journal of Southwest Jiaotong University* 41 (5): 571–574.
5. Si, Bingfeng, Baohua, Mao, and Zhili, Liu. 2007. Passenger flow assignment model and algorithm for urban railway traffic network under the condition of seamless transfer. *Journal of the China Railway Society* 29 (6): 12–18.
6. Xu, Ruihua, Qin, Luo, and Peng, Gao. 2009. Passenger flow distribution model and algorithm for urban rail transit network based on multi-route choice. *Journal of the China Railway Society* 31 (2): 110–114.

7. Zhang, Tianran. 2010. Optimal strategy for rail transit assignment model improvement. *Urban Mass Transit* 13 (9): 48–51, 56.
8. Song, Yaping, Ruijun, Wang, and Huiling, Li. 2011. Optimization model and algorithm of passenger flow assignment based on comprehensive transportation network. *China Railway Science* 32 (5): 127–131.
9. Lu, Limeng, and Jianping, Liu. 2011. Passenger flow distribution in rail transit no-barrier transfer mode. *Journal of Computer Applications* 31 (11): 3126–128.7.
10. Lin, Zhan, Mingqing, Jiang, Jianfeng, Liu, and Si, Bingfeng. 2012. Improved Logit model and method for urban rail transit network assignment. *Journal of Transportation Systems Engineering and Information Technology* 12 (6): 145–151.
11. Zhang, Tianran. Dongyuan, Yang, and Yali, Zhao. 2008. Comparative study of RP/SP combined data estimation between mixed logit and nested logit model. *Journal of Tongji University (Natural Science)* 36 (8): 1073–1078, 1084.
12. Xie, Ruhe, Zhuqiang, Qiu, and Qingyun, Li. 2006. Application of Logit model in estimating the distribution rate of passenger flows on Guangzhou–Shenzhen Railway. *China Railway Science* 27 (3): 111–115.

Path Selection Research with Digestion Index

Xiaoyu Meng, Hao Yue and Xiaoling Liu

Abstract To study the path selection's influence on the network's transmission and processing capability, we describe the corresponding scale-free network's traffic dynamics and analyze the final network states under different packet generation rates. Also, we introduce the digestion index representing the network's transmission and processing capability. With the digestion index of different paths, the new route of aviation network can choose the combined path.

Keywords Digestion index · Path selection · Scale-free network · Aviation network

Introduction

Almost all of networks in real life can be abstracted as complex networks of nodes and edges. If the city airport as nodes, these nodes are connected by air routes, which make up urban air network. If Web pages as nodes, hyperlinks as connections between pages, they constitute the Internet communication network. With the theoretical research gradually deepening, the topological characteristics of urban traffic network based on complex network have become the research focus of many researchers [1–3]. A large number of empirical studies have found some traffic networks with scale-free properties. Sen et al. [4] studied the Indian railroad, and found out its small-world characteristic. Latora and Marchiori [5] studied preliminarily the Boston subway's small-world characteristic. Guimera and Amaral [6] analyzed the statistic characteristics of the worldwide aviation network and concluded that the network had an obvious scale-free property. Wu et al. [7] carried out the complex network theories' empirical research and studied the scale-free property of Beijing bus network. Gao and Li [8] studied the scale-free traffic network

X. Meng · H. Yue (✉) · X. Liu
MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology,
Beijing Jiaotong University, Beijing, China
e-mail: 14120867@bjtu.edu.cn

and revealed its internal evolutionary mechanism. Boccaletti et al. [9] reviewed that the evolution process of some traffic network showed the scale-free properties. In this paper, we study in detail path selection behaviors of scale-free network. It will effectively improve network's information transmission and processing capability, thereby providing the theoretical foundation for guiding practice.

Network Description

The network is composed of nodes and edges. The node's degree refers to the total number of edges through a node and that is the number of other nodes (called neighbors) that have connections with the node. The average value of all nodes' degree is called the network's average degree, denoted $<K>$.

The shortest path, called the distance between two nodes, refers to the minimum number of passed nodes connecting two nodes in the network. The average value of all node pairs' distance is called the network's average path length.

With the improvement of computers' running ability, it is possible to collect real network systems' data and analyze the topology structure characteristics, which greatly promote the development of the complex networks. Table 1 shows some real networks' basic statistical data [10], where N is the number of all nodes, M is the number of all edges, $<K>$ is the network's average degree, and L is the network's average path length.

Traffic Dynamic

The algorithm of the scale-free network model is as follows: [11, 12].

To embody the growing character of the network, starting with a few number of vertices (m_0), at every time step we add a new vertex with m ($\leq m_0$) edges that link the new vertex to m different vertices already present in the network.

To embody preferential attachment, we assume that the probability P that a new vertex will be connected to the present vertex i depends on the degree k_i of the vertex i, such that

Table 1 Some feature parameters of empirical networks

Network type	N	M	$<K>$	L
WWW (nd)	269,504	1,497,135	5.55	11.3
WWW (Alta)	2.04E+08	2.13E+09	10.5	16.2
Internet (AS)	10,697	31,992	5.98	3.31
Chinese air	121	1378	11.4	2.22
American air	215	116,725	272	2.4
Indian air	79	442	5.59	2.26
Global air	3883	18,810	4.84	4.37

$$\prod(k_i) = \frac{k_i}{\sum_j k_j} \tag{1}$$

where k_i is the degree of the new vertex i, and j refers to the vertex already present in the network.

There are a few nodes called the hub nodes, each of which has an enormous degree in the scale-free network. With complex network model, we can study the transmission of traffic information flow based on the traffic dynamics through computer simulations. The transmission media and information in real network, such as Internet data packet and aviation aircraft, can be abstracted from the network model's mobile units, called "packets." The rules of traffic dynamic are as follows:

At every time step we add R packets into the network, each of which selects origin node and destination node randomly. R is called the packet generation rate.

The maximum number of packets a node can transmit is called the node's processing ability, denoted C. The packets are sent in accordance with the first in first out (FIFO) rule.

At every time step, each packet can at most moved to next node from current node. If the packet arrives at the destination node, the packet is removed from the network.

According to the different R values, the network's final states are divided into steady state and hectic state.

Steady state means that the number of the newly added packets is almost equal to the number of removed packets. So the total number of packets in the network remains almost unchanged.

Hectic state means that a fraction of the added packets can reach their destination nodes and be removed from the network, but most of the packets are accumulated in the network. With the total number of packets in the network increasing, the network shows the hectic state.

The phase transition from the steady state to the hectic state is the key to the network's transmission and processing capability. The generation rate R_c corresponding to the phase transition is an important indicator, called the critical generation rate.

Path Selection

First, we set up the scale-free network, and structure different sequence of paths between nodes, which can be achieved through C++ program. Next, we simulate the packets' transmission based on the present path through the computer, as shown in Fig. 1. Finally, we change the value of R to observe the final state of the network and get the value of R_c.
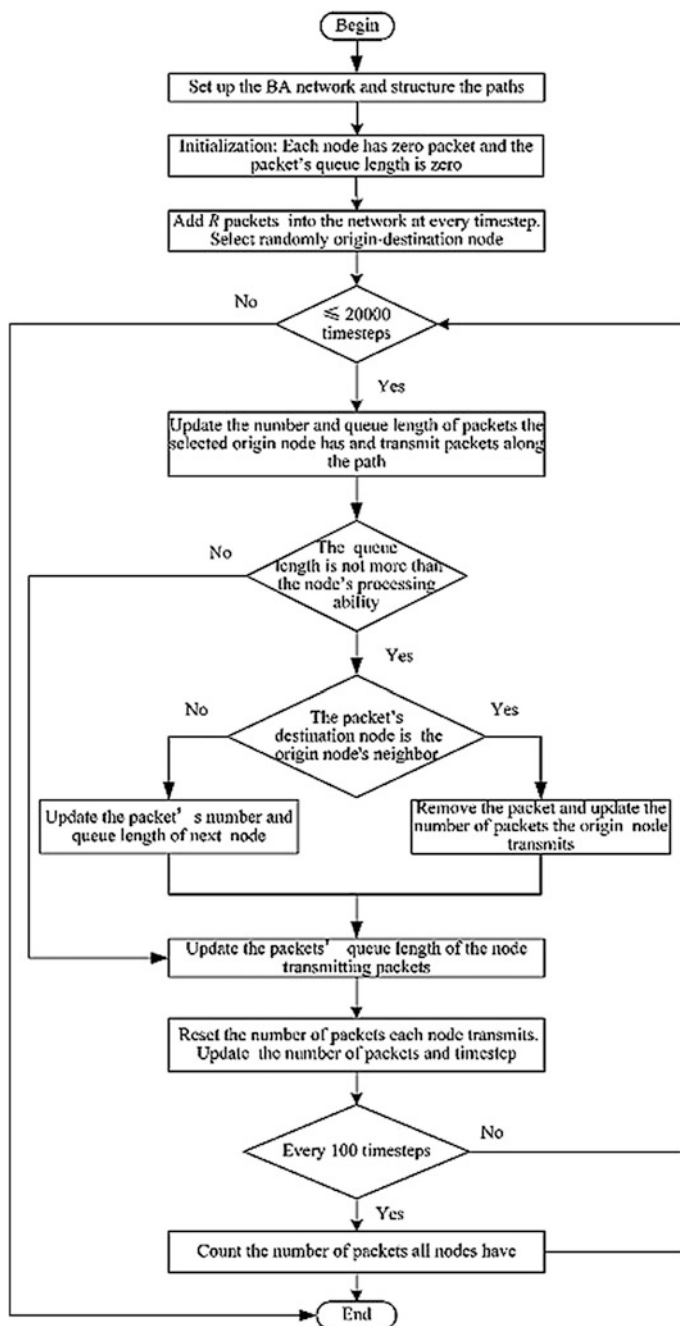
Fig. 1 Packets' transmission with the path
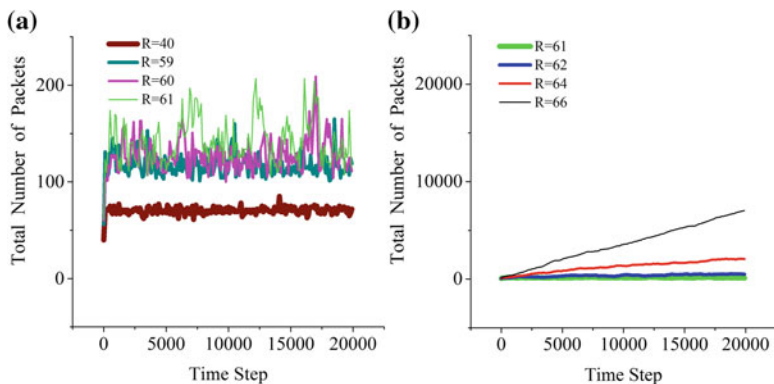
Fig. 2 Packets with the shortest path

The shortest path means that all nodes in the network are at the same level of importance, and the differences of the nodes' degree are ignored, which cause the path failing to make full use of the nodes that have good accessibility. Also, the transmitted packets are easy to be in the queue and stranded at the nodes that have bigger degrees, which aggravate the nodes' burdens and influence the network's information transmission capability. Taking into account the incompleteness of the shortest path, we propose a new path, that is, the minimum sum of passed nodes' degree in the transmission process of the packet from the origin node to the destination node. In this way, the transmitted packets can bypass the node that has bigger degree where a lot of packets come together and select the node that has fewer degree and better accessibility so that the new path can improve network's information transmission capability.

Without loss of generality, we analyze the packets' transmission process taking $N = 1000$, $C = 3$, and $<K> = 14$ as an example. As Fig. 2 demonstrates, the total number of packets in the network tends to a stable value when $R = 40$. The network shows steady state. When $R = 66$, the total number of packets in the network increases dramatically. The network shows hectic state. Furthermore, when $R \leq 61$, the network shows steady state. When $R > 61$, the network shows hectic state. So the critical packet generation rate R_c^s is equal to 61 with the shortest path. Similarly, as Fig. 3 demonstrates, the critical packet generation rate R_c^n is up to 342 with the new path. That is consistent with the fact. The new path means that the fights give preference to unimpeded routes, reduce delays at the hectic routes, and improve resource utilization. Therefore, the new path is conducive to the shunt of routes and enhances the network's transmission and processing capability. The digestion index of the network is as follows:

$$I = R/P \tag{2}$$

where R is the packet generation rate, and P refers to the total number of packets in the network.
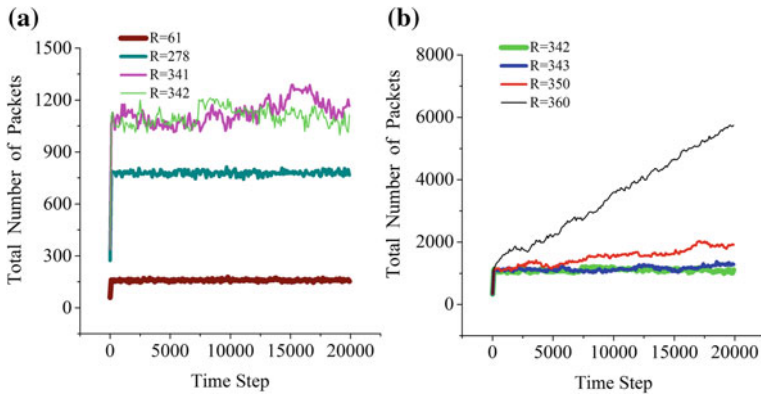
Fig. 3 Packets with the new path

Table 2 Digestion index with the shortest path

R	10	20	30	40	50	60	61	62	70
P	17	35	53	71	91	125	152	290	7334
I	0.588	0.571	0.566	0.563	0.549	0.480	0.401	0.214	0.010

Table 3 Digestion index with the new path

R	10	20	30	40	50	60	61	62	70
P	26	53	78	105	131	157	159	162	183
I	0.385	0.377	0.385	0.381	0.382	0.382	0.384	0.383	0.382
R	100	200	300	342	343	350			
P	263	537	863	1101	1111	1558			
I	0.380	0.372	0.348	0.311	0.309	0.225			

As the above shows, the value of I increases with the increase of R and the reduction of P, which means the increase of the network's transmission and processing capability. It should be pointed out that the value of P refers to the average value with all time steps for the steady state, but that is the median for the hectic state to avoid the extremum's influence. The value of I is shown in Table 2 with the shortest path and the new path's value is shown in Table 3.

As Tables 2 and 3 show, the digestion index I^s with the shortest path reduces significantly when $R = 62$, but the digestion index I^n with the new path correspondingly remains at an higher level. Besides, the value of I^s surpasses the value of I^n when $R \leq R_c^s$, but the new path has an advantage over the shortest path when $R_c^s \leq R \leq R_c^n$. The digestion index's trend with the path is shown in Fig. 4.

Fig. 4 Digestion index with the path
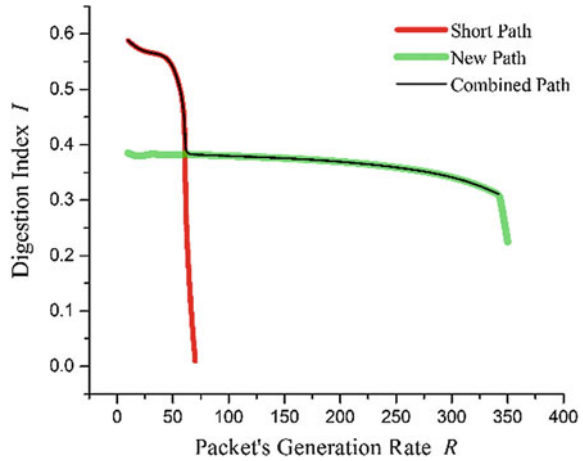
Conclusion

This paper presents the network's final state by the packet generation rate and study the network's transmission and processing capability with the different paths based on the digestion index. Accordingly, we propose the new route's choice approach of the aviation network under the steady state. Assuming that there have been p routes in the aviation network and there will be r new routes added, the ratio of r and p is crucial to choose the new routes. If the ratio is not less than the shortest path's critical digestion index I_c^s (corresponding to R_c^s), the new routes should choose the shortest path. If the ratio ranges from the shortest path's critical digestion index I_c^s to the new path's critical digestion index I_c^n (corresponding to R_c^n), the added routes should choose the new path. It is inappropriate to generalize about the route choice. We should comprehensively take into account the routes to give full play to their complementary advantages. It is further considered to set up the accurate model corresponding with the real network and extract deeply the traffic dynamic characteristics in future.

References

1. Tadic, B., and S. Thurner. 2005. Search and topology aspects in transport on scale-free networks. *Physica A: Statistical Mechanics and its Applications* 346 (1): 183–190.
2. Gao, Z.Y., et al. 2006. Research on problems related to complex networks and urban traffic systems. *Journal of Transportation Systems Engineering and Information Technology* 6 (3): 41–47.

3. Cardillo, A., et al. 2006. Structural properties of planar graph of urban street patterns. *Physical Review E* 73 (6): 066107.
4. Sen, P., et al. 2003. Small-world properties of the Indian Railway network. *Physical Review E* 67 (3): 036106.
5. Latora, V., and M. Marchiori. 2003. Economic small-world behavior in weighted networks. *European Physical Journal B: Condensed Matter and Complex Systems* 32 (2): 249–263.
6. Guimera, R., and L. Amaral. 2004. Modeling the world-wide airport network. *European Physical Journal B* 38 (2): 381–385.
7. Wu, J.J., et al. 2004. Urban transit as a scale-free network. *Modern Physics Letters B* 18 (19): 1043–1049.
8. Gao, Z.Y., and K.P. Li. 2005. Evolution of traffic flow with scale-free topology. *Chinese Physics Letter* 22 (10): 2711–2714.
9. Boccaletti, S., et al. 2006. Complex networks: Structure and dynamics. *Physics Reports* 424 (4): 175–308.
10. Newman, M.E.J. 2002. The structure and function of networks. *Computer Physics Communications* 147 (1): 40–45.
11. Barabási, A.L., and R. Albert. 1999. Emergence of scaling in random networks. *Science* 286 (5439): 509–512.
12. Wang, X.F., et al. 2006. *Complex networks: Theory and application*. Beijing: Tsinghua University Press.

Dynamic Timetables Optimization Method of Regional Public Transit Under APTS

Xiqiao Zhang, Zhichao Sun and Wen Cui

Abstract The core of real-time operational control of public transit is compiling dynamic timetables. Considering the practical need of transfer time, this paper designs a new strategy for developing public transit timetables based on transfer time window: Firstly, for single transfer stations, aiming at minimizing the waiting time of passengers both in buses and outside buses, stop orders and dwell time are designed based on the degree of importance of transfer. Secondly, optimization model of regional public transit timetables is established and the model is the basis for optimization of public transit timetables. Thirdly, the model is solved by using genetic algorithm. Finally, practical case is given to verify the performance of the model.

Keywords Advanced public transportation system · Regional public transit · Dynamic timetables · Genetic algorithm

Introduction

With the rapid development of advanced public transportation system (APTS), technology related to information collection and publication has become more and more mature, which provides foundations for the investigation of real-time operational control of public transit. The core of real-time operational control of public transit is compiling dynamic timetables, i.e., compiling a timetable that makes buses from different lines arrive at certain transfer station simultaneously on condition that we have known the real-time information of passengers, buses and transfer, so that we can minimize passengers waiting time, improve the level of service and operational efficiency, and enhance the attractiveness of public transit.

X. Zhang · Z. Sun (✉) · W. Cui
School of Transportation Science and Engineering,
Harbin Institute of Technology, Harbin, China
e-mail: hitszc@163.com

Starting from the 1970s, scholars began to study regional public transit timetables which mainly focus on the coordination and optimization of transit timetables in transportation system. In 1976, Rapp and Gehner [1] described a graphical interactive tool which optimizes the regional public transit system through the way of adjusting the departure frequency of buses under the constraint of operating cost. In 1981, Furth and Wilson [2] established the optimization model aiming at minimizing the total waiting time and maximizing the social benefit under the constraints of total cost, rated passenger capacity and rated load, determined regional transit departure frequency. In 1984, Ceder [3] presented four different ways to determine departure interval which become the basis for subsequent researches. Researches on public transit timetables in recent years mainly aim at setting up mathematical models which consider the knowledge of relevant subjects, solving the model through ways of various optimization algorithms, and programming to determine the departure frequency and departure interval. In 2001, Ceder [4] established a mixed-integer programming model (MILP) which aims at maximizing the number of buses from different lines arriving at the same transfer site; simultaneously, he solved the model by using Turbo-Pascal and applied it to multi-line collaboration [5–8]. In 2005, Fleurent [9] established a network flow model, the model introduced three types of transfer time (minimum, maximum, and ideal), formed composite sync-transfer quality indicators. In 2010, Hadas and Ceder [10] proposed reducing travel time to improve the quality of public transit service. They presented that the way to reducing travel time is coordinated public transit timetables, but this type of timetables is limited by the uncertainty of the number of stop buses. They established a dynamic programming model and got a series of optimization results. In 2012, based on the public transit network in Monterrey, Omar [11] established model aiming at maximize the meeting times for buses on different lines, he discovered a special structure of the solution domain of BTP issue, so that he reduced the number of decision variables and constraints by preprocessing. In 2013, Wu and Tang [12] proposed a new mathematical programming model for reliable timetables design problem of public transit network that considers the stochastic bus travel times. The new formulation simultaneously optimized the decision variables of the departure time of each trip for each line and the slack time associated with transfer station.

Most of the existing regional public transit timetables only try their utmost to realize the synchronous transfer and improve the level of transfer theoretically. But they ignore the minimum time needed for transfer. So that this paper designs a new strategy for developing public transit timetables based on transfer time window.

Strategy for Bus Scheduling Based on Dynamic Operation Timetables

Based on the technology of real-time operational control of public transit and dynamic operation timetables, this paper designs a strategy for bus scheduling which arranges the meeting time and dwell time of buses on different lines according to the demand for transfer: Firstly, bus dispatch center acquires the information of passengers' travel time, routes, transfer site, etc., by way of history travel information and the query information before trip. Secondly, the center forecasts the number of passengers who have demand for transfer and generates a reasonable initial scheduling strategy according to the number of buses being operating on network. Because of the randomness of real-time information, it is very hard to get the accurate mathematical simulation of public transit passenger flow, so that the APTS constantly modifies the meeting time and dwell time of buses during the implementation of timetables and finally push it to bus drivers through MsgExperss and realizes dynamic operation timetables. After above process, we can minimize passengers waiting time, improving the level of service and operational efficiency, and enhancing the attractiveness of public transit.

In the process of scheduling, we use automatic vehicle location (AVL) to realize the real-time control of stop orders and dwell time which achieves efficient use of time and space. The model will screen out the buses on the lines which have greater degree of importance of transfer (i.e., buses contain a larger number of transferring passengers), let them arrive at the transfer site in advance. There are three steps for passengers arriving at transfer site in advance to complete the transfer; besides getting on and off the buses, there is a time interval between getting on and off; in this paper, we define it as transfer time window. Explicit process is shown in Fig. 1a, b. The reasonableness of the setting of time window directly affects the effect of dynamic timetables, its size depends on the number of transfer passengers on different lines at certain transfer station.
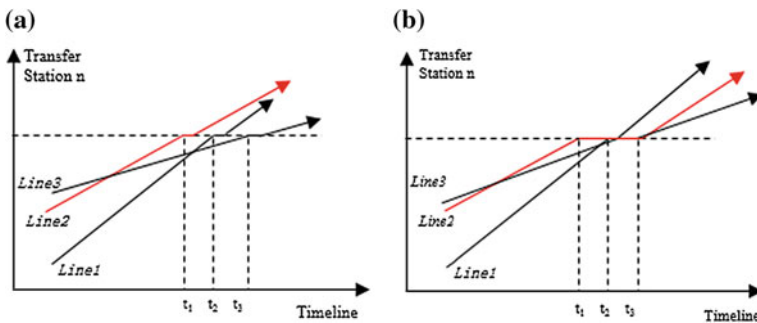
Fig. 1 **a** Before optimization. **b** After optimization

Fig. 2 A simple transit network
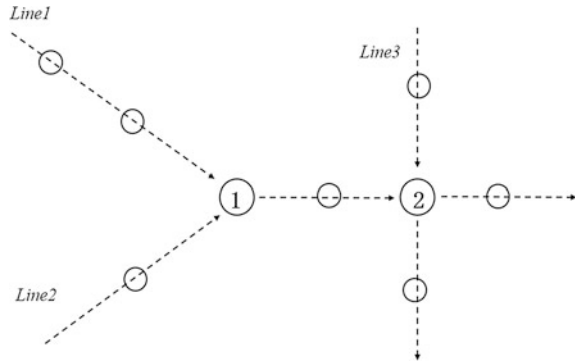
Defining the Model

Public transit network consists of bus lines and transfer stations; because this paper aims at optimizing public transit timetables on the basis of transfer time window, we just consider the transfer station. Fig. 2 shows a simple public transit network with 3 bus lines and 2 transfer stations. The collection of all bus lines is expressed as I, the collection of bus lines on which transfer happens is expressed as $J_n(i)$. Figure 3 shows the process of bus arrival, passengers transfer, and bus departure, the definition of the parameters in this paper are shown in Table 1.
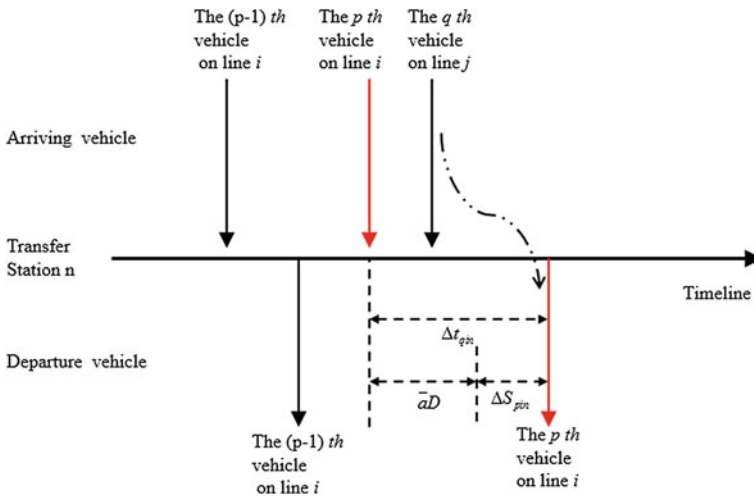
Fig. 3 Process of passengers transfer

Table 1 Definition of parameters in the model

Parameters	Table column head
M	The number of bus lines on network
N	The number of bus lines on network
F_i	The number of stations on line i
C_{jin}^q	The number of passengers transferring from line j to line i on bus q and station n
X_{pi}	Departure time of bus p on line i
H_i^M, H_i^m	The upper and lower limit of departure interval
T_{pin}	The operation time of bus p on line i from start to station n
d_{pin}	The time when bus p on line i arriving at station n
I_{pin}	The time when bus p on line i leaving station n
Δt_{pin}	The waiting time of bus p on line i and station n
R_{pin}	The number of passengers in bus p on line i when leaving station n
$\bar{a}$	The average time of getting on and getting off buses (min/per)
O_{pin}	The number of passengers getting on bus p on line i when arriving at station n
U_{pin}	The number of passengers getting off bus p on line i when arriving at station n
ΔS_{pin}	The transfer time window of bus p on line i when arriving at station n
ΔS^M	The upper limit of transfer time window
ξ	The limit of departure interval
ε	The shortest time of getting on and getting off buses
V_{qjn}	The weight of bus q on line j and station n
W_n	The weight of station n
α	The waiting time of transfer passengers
β	The waiting time of passengers in buses

Basic Assumption

The basic assumptions of the model in this paper are as follows:

1. The passenger flow based on APTS is known;
2. The impact of the road traffic conditions does not take into consideration;
3. The travel time between two adjacent stations is known;
4. The scheduling cycle is fixed (60 min);
5. The operation of all buses is strictly in accordance with the schedule, and excluding the case of midway turn;
6. The buses operating on the road are all considered as whole cycling i.e., all buses leave from the starting stations, reach the terminal after passing through all transfer stations on their lines.

Modeling

Table 1 shows the definition of parameters in the model.

1. The transfer time of all passengers on network is:

$$T_1 = W_n \sum_{n=1}^{N} \left(\sum_{i=1}^{M} \sum_{p=1}^{F_i} \sum_{j=1}^{M} \sum_{q=1}^{F_j} \left(V_{qjn} C_{jin}^q \left(I_{pin} - d_{qjn} \right) \right) \right) \tag{1}$$

2. Waiting time of passengers in buses is:

$$T_2 = \sum_{n=1}^{N} \sum_{i=1}^{M} \sum_{p=1}^{F_i} \Delta t_{pin} \left(R_{pi(n-1)} - O_{pin} \right) \tag{2}$$

The model for regional public transit schedule is as follows:
Objective function:

$$\min(\alpha T_1 + \beta T_2) \tag{3}$$

i.e.,

$$\min \left(\begin{array}{l} \alpha W_n \sum_{n=1}^{N} \left(\sum_{i=1}^{M} \sum_{p=1}^{F_i} \sum_{j=1}^{M} \sum_{q=1}^{F_j} \left(V_{qjn} C_{jin}^q \left(I_{pin} - d_{qjn} \right) \right) \right) \\ + \beta T_2 = \sum_{n=1}^{N} \sum_{i=1}^{M} \sum_{p=1}^{F_i} \Delta t_{pin} \left(R_{pi(n-1)} - O_{pin} \right) \end{array} \right)$$

Constraints:

$$X_{1i} \le H_i^M \tag{4}$$

$$T - H_i^M \le X_{F_i i} \le T \tag{5}$$

$$H_i^m \le X_{(p+1)i} - X_{pi} \le H_i^M \tag{6}$$

$$\left| \left(X_{(p+1)i} - X_{pi} \right) - \left(X_{pi} - X_{(p-1)i} \right) \right| \le \xi \tag{7}$$

$$I_{(p-1)in} < d_{qjn} + \varepsilon \le I_{pin} \tag{8}$$

$$0 \leq \Delta S_{pin} \leq \Delta S^M \tag{9}$$

Constraint (4) ensures the departure time of the 1st bus complies with the requirement. Constraints (5) ensure the departure time of the last bus does in the range of plan. Constraint (6) indicates the minimum and maximum departure time interval limits. Constraints (7) indicate the limit value of time interval between two adjacent buses. Constraint (8) ensures that passengers have sufficient time achieving the process of transfer. Constraint (9) ensures the maximum value of transfer time window.

Parameter Calibration

Determining the Weight of Lines

The weight of lines represents the contribution of different lines to the reduction of waiting time on the network, so that it determines bus stop orders and the time interval between two stops. Because the model in this paper takes the practical transfer of passengers into account, we select the flow of getting on buses, getting off buses, and flow of transfer passengers as a parameter assessing the weight of lines:

$$V_{qjn} = \begin{cases} 0 & , C_{jin}^q = 0 \\ \dfrac{\max\left(U_{qjn} \times \bar{a}, O_{qjn} \times \bar{a}, \sum_{i=1}^{M} C_{jin}^q + \varepsilon\right)}{\sum_{j=1}^{M} \sum_{q=1}^{F_i} \max\left(U_{qjn} \times \bar{a}, O_{qjn} \times \bar{a}, \sum_{i=1}^{M} C_{jin}^q + \varepsilon\right)} & , C_{jin}^q \neq 0 \end{cases}$$

Determining the Weight of Transfer Stations

The weight of transfer stations represents the contribution of transfer stations to the reduction of waiting time on the network. This paper chooses the number of transfer passengers on certain transfer station as a parameter assessing the weight of transfer stations, where the number of transfer passengers is the number of passengers achieving transfer on certain station during the plan period:

$$W_n = \frac{\sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{q=1}^{F_j} C_{jin}^q}{\sum_{n=1}^{N} \sum_{i=1}^{M} \sum_{j=1}^{M} \sum_{q=1}^{F_j} C_{jin}^q} \tag{10}$$

Determining the Weight of Various Kinds of Waiting Time

Based on investigation, this paper uses Delphi method to determine the weight of the waiting time of transfer passengers and the waiting time of passengers in buses, the result is as follows:

$$\alpha = 2 \quad \lambda = 1$$

Determining the Upper and Lower Limits of Variables

1. After analysis of the data of getting on and getting off buses in different transfer stations, this paper determines the average time of getting on and getting off buses is 2 s i.e., $\bar{a} = 2$ s.
2. The theoretical upper limit of departure interval on line i is a follows:

$$H_i^M = \frac{60v_i}{Q_i} \tag{11}$$

where Q_i is the total number of passengers in planning year, v_i is the rated capacity of the buses on line i.

3. Usually, the lower limit of departure interval is no less than 1–3 min, so that this paper determines the theoretical lower limit of departure interval as 3 min i.e., $H_i^M = 3$.
4. In order to ensure the continuity of the departure, this paper determines the difference between two adjacent departures intervals is 3 min i.e., $\xi = 3$.
5. The shortest time of getting on and getting off buses is 10 s, i.e., $\varepsilon = 10$.
6. When determining the value of transfer time window, we should minimize the delay caused by transfer. According to investigation, the acceptable waiting time of passengers is 60 s. As determined by this paper, the upper limit of the value of transfer time window is 30 s, i.e., $\Delta S^M = 30$.

Solving Model

This paper solves the model by using the genetic algorithm. The basic idea is generating the original programs, then, adjusting the program by using the genetic algorithm to obtain the optimal solution to the model, finally, getting the most
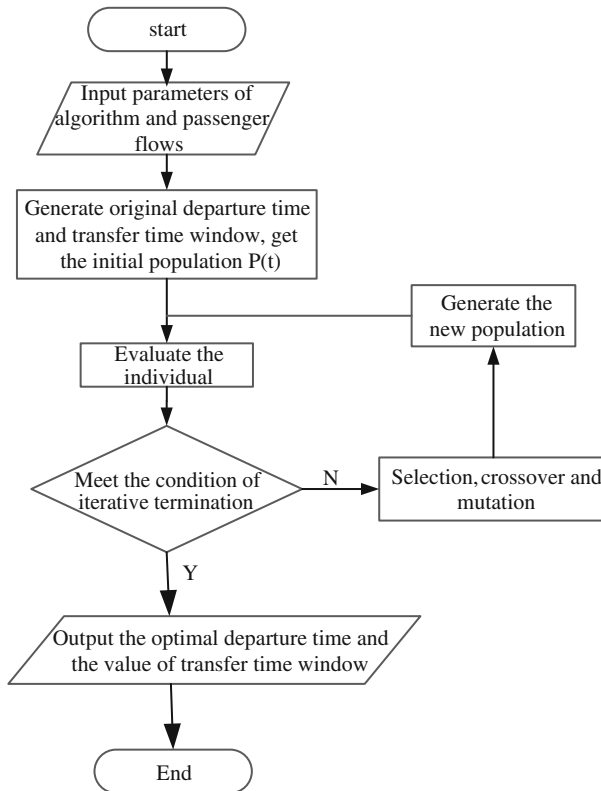
Fig. 4 Process of passengers transfer

reasonable way of bus schedules to help transit managers make the right decisions. Specific steps are shown in Fig. 4.

Scholars have previously used genetic algorithms to solve the scheduling problem of single line. The program in this paper is based on the design of algorithms from previous research results and combines the characteristics of regional bus timetables generation model. Specific operating steps are as follows:

1. Encoding

There are two decision variables in this model, the first is all trips departure time on the network X_{pi}, the second is the value of transfer window of all trips in each transfer stations ΔS_{pin}. This paper uses the binary encoding to encode variables.

2. Initial population

First, the program generates a certain number of individuals randomly, then picks out the best individual, and adds it to the initial group; this process is iterated until the number of individuals in the initial population reaches the predetermined

scale. During the generation of initial group, all of the individuals will be tested, the program will retain the individuals which satisfy the constraints and eliminates the individuals which does not satisfy the constraints. The size of initial population is determined by the size of the network, generally, a population with 200 individuals is suitable for medium bus network. Population scale should be greater than or equal to the number of decision variables, so that each individual in the population exceeds the searching scope.

3. Fitness function

The process of fitness calculation is as follows: Firstly, decoding the individuals and obtaining the individual's phenotype X. Secondly, calculating the value of the objective function according to the phenotypes. Finally, according to the type of problems, calculating individual's fitness in accordance with certain rules by the value of objective function $Fit(f(x))$.

This paper uses the following rules to do the transformation:

$$Fit(f(x)) = \begin{cases} C_{max} - f(x), & C_{max} - f(x) > 0 \\ 0, & \text{else} \end{cases} \qquad (13)$$

where $f(x)$ is the value of objective function, C_{max} is the maximum estimated value for $f(x)$.

4. Selection

The arithmetic operators using the method called fitness ratio (Roulette Wheel Selection) specifies the field of the selection function as Roulette.m in the text box 'Selection.' To improve the quality of selection, the paper combines copying with Roulette Wheel Selection. By setting the number of best individual (Elite count) in text box 'Reproduction,' we can specify the number of replicate individual in the population.

5. Crossover

The paper selects the method called discrete recombination. That is, specifying the crossover function as Scattered.m in the text box 'Crossover.' Crossover fraction controls the crossover frequency, it may impact the stability of the system if the crossover fraction is too large or too small, generally, crossover rate is 0.25–0.75. We can specify the crossover rate by way of setting 'crossover fraction' in the text box 'Reproduction.'

6. Mutation

The program chooses an alterable mutation probability AA during the iteration process. In early stage, the program chooses lager AA to expand search space; while in late stage, the program chooses smaller AA to accelerate convergence. The recommended range of values is generally from 0.0001 to 0.2.

7. Termination conditions

This paper uses maximum iteration as a judging condition for termination, the number of iteration is determined by the scale of network, and the recommended range of values is generally from 10 to 600.

Example Analysis

Example Designing

Figure 5 shows a public transit network with 3 bus lines and 1 transfer station; line 1 and line 3 are both pass the transfer station. The planning operating parameters are shown in Table 2.

After the fitting of passengers' flow data, passenger arrival rate functions about the three lines are obtained, for example, arrival rate function about line 1 is:

$$f_{11}(t) = 0.002t^4 - 0.0061t^3 + 0.0826t^2$$
$$- 0.3859t + 1.2830$$

The getting off rate and transfer rate of line 1, 2, and 3 on each station are shown in Tables 3 and 4.

Result Analysis

This paper solves the example with the help of MATLAB genetic algorithm Toolbox. To validate model effect, this paper considers three types of timetables

Fig. 5 Sketch of public transit network
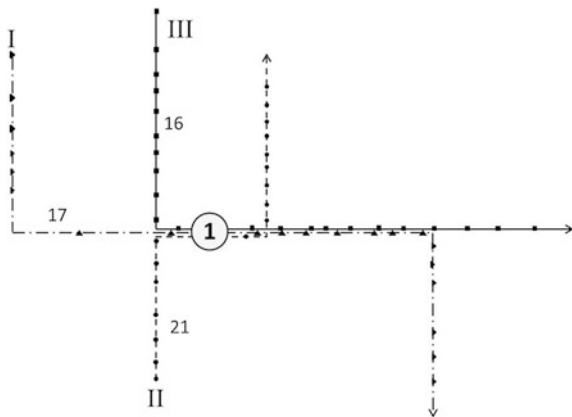
Table 2 Planning operating parameters of three lines

Lines	I (8)	II (7)	III (18)
Passenger flow during planning horizon (per/h)	560	672	336
Line length (km)	12.84	13.46	11.48
The number of stations	23	24	19
Bus rated capacity (per)	70	70	70
Upper limit of departure interval (min)	8	7	13
Lower limit of departure interval (min)	3	3	3
ID of transfer station ① on line	9	12	8

Table 3 Getting rate on each station (%)

Stations	1	2	3	4	5	6	7	8
I	0	0	0	3	20	30	18	30
II	0	0	0	5	15	20	20	40
III	0	0	0	0	20	30	30	20
Stations	9	10	11	12	13	14	15	16
I	20	20	50	40	20	35	28	15
II	30	40	20	60	50	5	40	25
III	20	50	40	10	60	10	20	0
Stations	17	18	19	20	21	22	23	24
I	50	38	60	40	20	55	100	–
II	20	30	50	40	20	30	60	100
III	40	50	100	–	–	–	–	–

Table 4 Transfer rate on each station (%)

Transfer rate	I	II	III
I	0	15	20
II	10	0	10
III	15	20	0

Table 5 Optimal result

Timetable type	Traditional one-way	Traditional regional	Dynamic timetable based on transfer time window
Passengers' waiting time	7993	6964	6650

(traditional one-way timetable, traditional regional timetable, and dynamic timetable based on transfer time window). The three cases' minimum waiting time of passengers on network is shown in Table 5.

As is shown in Table 5, on one hand, when using traditional regional timetables, the passengers' waiting time reduces 12.9% compared with traditional one-way timetables which proves that collaboration is beneficial to the reduction of waiting time i.e., it can improve the level of public transit service. On the other hand, when

Table 6 Optimal result

Bus ID	Before optimization	After optimization	Transfer time window (min)
Line I			
1	7:00	7:02	2
2	7:06	7:10	1
3	7:12	7:16	1
4	7:18	7:22	1
5	7:24	7:27	1
6	7:30	7:31	2
7	7:36	7:36	1
8	7:42	7:40	2
9	7:48	7:46	0
10	7:54	7:52	0
11	8:00	8:00	0
Line II			
1	7:00	7:00	0
2	7:05	7:07	0
3	7:10	7:13	0
4	7:15	7:19	0
5	7:20	7:24	0
6	7:25	7:29	0
7	7:30	7:33	0
8	7:35	7:36	2
9	7:40	7:40	2
10	7:45	7:45	0
11	7:50	7:51	0
12	7:55	7:55	1
13	8:00	7:59	0
Line III			
1	7:00	7:04	1
2	7:10	7:16	2
3	7:20	7:26	0
4	7:30	7:34	0
5	7:40	7:43	0
6	7:50	7:53	0
7	7:60	8:00	1

using dynamic timetables based on transfer time window, the passengers' waiting time reduces 4.5% compared with traditional regional timetables, so that we can conclude that dynamic timetables based on transfer time window is the most effective strategy to improve the level of public transit service.

The details of setting of dynamic timetables based on transfer time window are shown in Table 6.

Conclusion

Based on APTS, this paper studies the regional public transit schedule problem under the assumption that we have known the real-time information of passengers and buses:

1. To increase the probability of buses' 'encounter,' this paper designs a new strategy for developing public transit timetables based on transfer time window, and coordinates the regional public transit timetables with transfer time window.
2. Considering the practical need of transfer time, this paper establishes the model aiming at minimizing the waiting time of passengers both in buses and outside buses under different case of departure intervals.
3. The model is solved by using genetic algorithm which makes it possible to get satisfactory solutions within reasonable time limit.
4. The further studies would need to realize the connection of transfer time window, timetables, and real-time public transit demand.

References

1. Rapp, M.H., and C. Gehner. 1976. Transfer optimization in an interactive graphic system for transit planning. *Transportation Research Record* 619: 27–33.
2. Furth, P., and N. Wilson. 1981. Setting frequencies on bus routes: Theory and practice. *Transportation Research Record* 818: 17–22.
3. Ceder, A. 1984. Bus frequency determination using passenger count data. *Transportation Research Part A: General* 5: 439–453.
4. Ceder, A., B. Golany, and O. Tal. 2001. Creating bus timetables with maximal synchronization. *Transportation Research Part A: Policy and Practice* 10: 913–928.
5. Ceder, A. 2001. Bus timetables with even passenger loads as opposed to even headways. *Transportation Research Record: Journal of the Transportation Research Board* 1760 (1): 3–9.
6. Ceder, A. 2001. Creating bus timetables with maximal synchronization. *Transportation Research Part A: Policy and Practice* 35 (10): 913–928.
7. Ceder, A., and H.I. Stern. 1981. Deficit function bus scheduling with deadheading trip insertions for fleet size reduction. *Transportation Science* 15 (4): 338–363.
8. Ceder, A., and O. Tal. 2001. Designing synchronization into bus timetables. *Transportation Research Record Journal of the Transportation Research Board* 1760 (1): 28–33.
9. Fleurent, C., and R. Lessard. 2005. Transit timetable synchronization: Evaluation and optimization. In *Proceedings of the 9th International Conference on Computer Aided Scheduling in Public Transport (CASPT), San Diego, California, Aug, 2005.*

10. Hadas, Y., and A. Ceder. 2010. Optimal coordination of public-transit vehicles using operational tactics examined by simulation. *Transportation Research Part C: Emerging Technologies* 6: 879–895.
11. Ibarra-rojas, O. J., and Y. A. Rios-solis. 2012. Synchronization of bus timetabling. *Transportation Research Part B: Emerging Technologies* 46: 599–614.
12. Wu, Y., and J. Tang. 2013. Optimizing reliable timetable for bus transit network: Model formulation and solution. In *25th Chinese Control and Decision Conference (CCDC), China*, 1835–1840.

Author Biographies

Xiqiao Zhang received the B.E. degree in traffic engineering, the M.E. degree in transportation planning and management, and Ph.D. degree in bridge and tunnel engineering from Harbin Institute of Technology, Harbin, China, in 2001, 2004, and 2008, respectively. He is currently an Assistant Professor with the Department of Transportation Engineering, School of Transportation Science and Technology, Harbin Institute of Technology. His research interests include traffic data analysis, transportation planning, and emergency traffic modeling and management.

Zhichao Sun received the B.E. degree in traffic engineering from Harbin Institute of Technology at Weihai, Weihai, China, in 2014. He is currently working toward the M.E. in transportation planning and management, Harbin Institute of Technology, Harbin, China. His research interests include intelligent transportation systems and transportation planning.

Wen Cui received the B.E. degree in traffic engineering from Harbin Institute of Technology, Harbin, China, in 2013. He is currently working toward the M.E. degree with the Department of Transportation Engineering, School of Transportation Science and Technology, Harbin Institute of Technology.

Study on the Low-Carbon Operating Evaluation Model in Expressway Rest Area

Shu Wang, Jianyou Zhao, Kaifang Wang and Yang Liu

Abstract This paper mainly studies evaluation of low-carbonization operation in freeway rest area. Taking function of service area as starting point and perspective of energy conservation and emission reduction and the increase of carbon sinks, the thesis first builds the structure of evaluation of low-carbonization operation in freeway rest area while expresses index connotation and its quantitative method. Furthermore, combining AHP and TOPSIS, the evaluation index and methods mentioned in the thesis are applied and tested by specific cases. The research shows that evaluation model proposed in the thesis is credible so that it needs more clean energy development and lighting system improvement in evaluation of low-carbonization operation in freeway rest area.

Introduction

The operation and management of service area is the key to offering potential value which may increase carbon sink and decrease carbon emission. Energy conversation in expressway service area does not only occur with service level improvement, it can also be provoked by the way the vehicles interact with each other and with the demand situation.

S. Wang
School of Economics and Management, Chang'an University,
Xi'an 710054, China
e-mail: 1285219150@qq.com

S. Wang · K. Wang
School of Economy and Trade, He'nan University of Technology,
Zheng'zhou 450001, He'nan, China

J. Zhao (✉) · Y. Liu
School of Automobile, Chang'an University, Xi'an 710064, Shaanxi, China
e-mail: jyzhao@chd.edu.cn

Y. Liu
e-mail: 544485894@qq.com

Wang [1] discussed operations management of expressway service area from the perspective of sustainable development. The ecological management is the key to the sustainable development of the service area. Du Yulin pointed out that the low carbon of service area is the new idea of design and construction. Wang Cheng [2] sums up the green evaluation indexes of expressway service area with the green building evaluation theory. Based on AHP, green evaluation system is established in the whole life cycle of expressway rest area. Ting-ting hou [3] verified that the expressway service area of artificial wetland system has better practicability and popularization value by ecological value method. Song Ruqiong built sustainable development evaluation model of service area, which should be coordinated with the regional society, economy, and environment. And combined with examples, it is verified that the model is feasible. Han Baoqing argued that energy conservation and environmental protection should be the necessary equipment of service area.

The above literatures for the construction of green, ecological service area provide the useful enlightenment. Based on above, the paper designs to increase carbon sink and reduces emission in expressway service area for energy saving, building the evaluation system of low-carbon operation.

The Connotation of Low-Carbonization Management of Expressway Service Area

Expressway Service Area

Expressway service area is an expressway facility, which is established to meet vehicle safety operation requirements and basic physiological and psychological needs of passengers during long time, closed driving on the expressway. The expressway service area should have the basic functions of public toilets, convenience stores, catering, accommodation, parking, refueling, car repair, etc. With the development of economy and society, the service area should not only meet the basic function of service guarantee but also adapt to the increasing social demand. And the extension of its service function is the inevitable requirement of social development. In addition to the basic functions, the service area should have multilevel services including specialty restaurants, supermarket scale, leisure and entertainment, information services, and extension functions including logistics, warehousing, cargo transit, transit passengers, and tourist service (Fig. 1).

Low-Carbonization Management of Expressway Service Area

The daily operations of the service area refer to a series of management activities after construction [1]. Its essence is to achieve the basic functions and extension
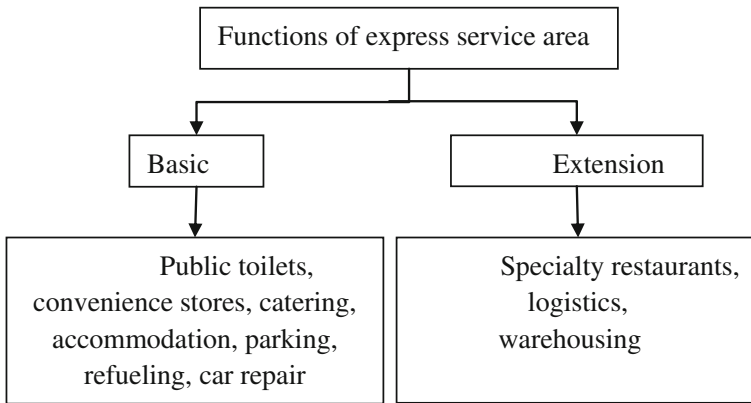
Fig. 1 Functions of express service area

functions of the service area through the provided water supply, electricity, equipment and maintenance services, and garbage and sewage treatment business. The function of expressway service area is accompanied with water, electricity, gas, and other energy consumption as well as treatment of domestic garbage and sewage, moreover, related to the service area of soil and water conservation, vegetation protection, etc., operation level is evaluated from two aspects of energy saving.

Therefore, in this paper, expressway service area low carbonization emission reduction and increase carbon sink. Energy saving and emission reduction refers to using ways and methods to reduce per unit of carbon emissions in daily operations of the service area, including the use of energy-saving technologies and equipment, use of clean energy, recycling of water resources utilization, waste of resources, and low-carbon management; increasing carbon sequestration refers to absorbing carbon dioxide in the air by vegetation, and treating domestic wastewater and degradable garbage by using of artificial wetland or aquatic ecological in daily operations of the service area.

The Low-Carbon Operation Evaluation System of Expressway Service Area

The Evaluation Framework Building of Low Carbon Operated in SERVICE Area

Depending on the principle of analytic hierarchy process the part takes on, the evaluation framework building of low carbon operated in service area is built. The evaluation framework contains six levels, as shown in Fig. 2.
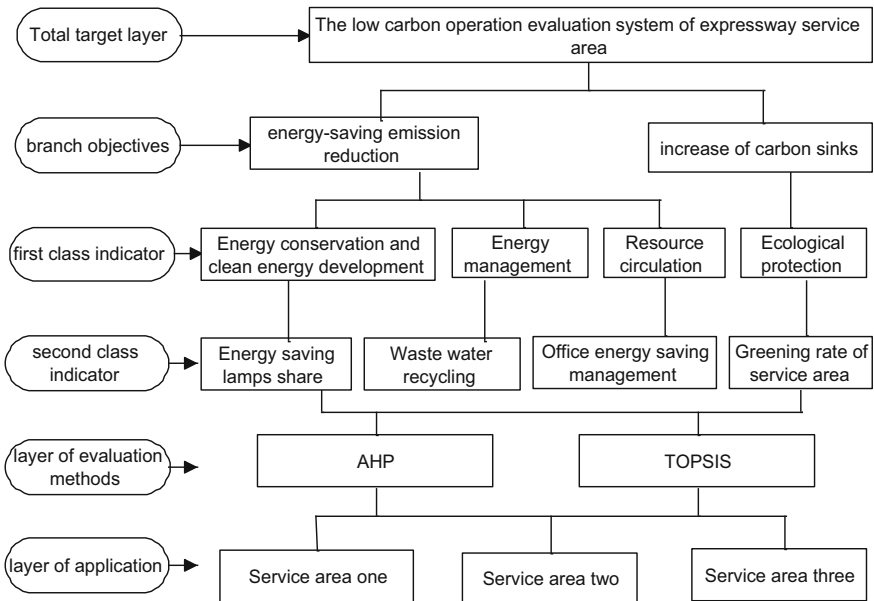
Fig. 2 Low-carbon operation evaluation system of expressway service area

The total target layer: The target of evaluation framework is to promote the expressway service area for low-carbon operation.

The target layer: According to the research angle of the article, the total target layer is broken into two branch objectives which are the energy-saving emission reduction and the increase of carbon sinks. The branch objectives of the energy-saving emission reduction are achieved by reducing the unit energy consumption as well as the developing clean energy; the increase of carbon sinks is mainly achieved by increasing vegetation and the ecological water.

Criterion layer: The criterion layer indicator of low carbon operated in service area includes four parts which are energy conservation and clean energy development, energy management, resource circulation, and ecological protection.

Index layer: Index layer involves 16 indicators. There are nine indicators from the related document and six indicators from personal setting.

The layer of evaluation methods: According to the characters of evaluation objects and survey data, the Analytic Hierarchy Process and TOPSIS method are used in the paper as the evaluation methods.

The layer of application: With the help of survey data from three services, the evaluation model this article takes on is applied and inspected.

Evaluation Index and Quantification of Service Area

As shown in Fig. 2, low-carbon operation evaluation system of the freeway service area includes four indicators of criterion layer, each first level index designs two indicators of some quantitative, its connotation and the quantitative methods are shown in Table 1.

1. Energy-saving and clean energy development (B_1). The energy-saving level of daily operation in the service area is measured by four secondary indicators, such as the ratio of energy-saving lamps (C_1), the level of energy-saving control measures (C_2), the ratio of energy-efficient office equipment (C_3), the ratio of water-saving appliances (C_4), the level of water-saving irrigation (C_5). Clean energy development mainly is measured through the percentage of clean energy accounted for all energy consumption (C_6). Among them, the qualitative indicators are using expert scoring (it is divided into 0, 1, 3, 5 from high to low), and quantitative indicators are given according to the percentage.

2. Energy-saving management (B_2). The level of operations management in the service area is measured by two three-level indicators, such as the office energy-saving system (C_7), post responsibility system (C_8), and energy-saving training system (C_9). The above three indicators are qualitative indicators. Expert groups do the quantitative scoring for different levels according to the three levels of indicators, that is, whether the system exists, the refinement degree of the system and the degree of implementation.

3. Resource recycling (B_3). Resources recycling level in the service area is measured through three secondary indicators, such as the percentage of wastewater recycling accounted for the total amount of wastewater (C_{10}), the percentage of waste recycling accounted for total waste (C_{11}), and recycling and utilization degree of rainwater (C_{12}). The first two indicators are quantitative index, which can be derived directly according to the percentage. A group of experts do the quantitative scoring for recycling and utilization degree of rainwater according to the three levels of indicators, that is, whether there is rainwater collection device, whether the device is used to collect rainwater, utilization degree of collecting rainwater.

4. Ecological protection (B_4). Ecological protection level mainly in the service area is measured through four secondary indicators, such as vegetation protection measures of the service area (C_{13}), vegetation growth of the service area (C_{14}), service area greening rate (C_{15}), and ecological water technology (C_{16}). Greening rate of the service area can be directly calculated by the percentage. The expert groups do the quantitative scoring for vegetation protection measures of the service area according to whether there is protection measure in the service area and the implementation degree of protection measures. Experts do the quantitative scoring for vegetation growth of the service area according to whether the plant is drought and pest phenomenon. Experts do the quantitative for the ecological water technology based on number and the using level of technology.

Table 1 Evaluation indexes and content of low carbon in freeway service area

Criterion layer	Index layer	Quantization of index layer	Index property	Index source
Energy-saving and clean energy development (B_1)	(C_1) The ratio of energy-saving lamps	Energy-saving lamps/the total amount of lamps	Quantitative	Setting index by oneself
	(C_2) The level of energy-saving control measures	Sectional lighting; Control of brightness; automatic power off; and so on	Qualitative	Wang Cheng [2]
	(C_3) The ratio of energy-efficient office equipment	Energy-saving equipment/office equipment	Quantitative	Setting index
	(C_4) The ratio of water-saving appliances	Water-saving apparatus/water consuming apparatus	Quantitative	Li xiao qin [4]
	(C_5) The level of water-saving irrigation	Mode and degree of green irrigation	Qualitative	Wang Cheng [2]
	(C_6) The percentage of clean energy accounted for all energy consumption	Clean energy/the total amount of energy consumption	Quantitative	Song ru liang and so on
Energy-saving management (B_2)	(C_7) The office energy-saving system	Whether it exist; the degree of refinement; the degree of implementation	Qualitative	Zhang Chen [5]
	(C_8) Post responsibility system		Qualitative	Duzhi yang
	(C_9) Energy-saving training system		Qualitative	Chen qiu hua and so on [6]
Resource recycling (B_3)	(C_{10}) The percentage of wastewater recycling accounted for the total amount of waste water	Recycling wastewater/total amount of wastewater	Quantitative	Setting index
	(C_{11}) The percentage of waste recycling accounted for total waste	Recycling waste/total amount of waste	Quantitative	
	(C_{12}) Recycling and utilization degree of rainwater	Recycling rainwater and utilization degree	Qualitative	

Table 1 (continued)

Criterion layer	Index layer	Quantization of index layer	Index property	Index source
Ecological protection (B_4)	(C_{13}) Vegetation protection measures of the service area	Whether there is protection measures and the degree of implementation	Qualitative	Setting index
	(C_{14}) Vegetation growth of the service area	Whether the plant is drought and pest phenomenon; situation of vegetation growth	Qualitative	Du yulin
	(C_{15}) Service area greening rate	The area of vegetation/the total area of service area	Quantitative	Setting index
	(C_{16}) Ecological water technology	Number and the using level of technology	Qualitative	

Note the qualitative indicators are using expert scoring (it is divided into 0, 1, 3, 5 from high to low). Experts can choose the transition scores between the four grades, according to the content and the actual situation of different indicators. Quantitative indicators are given according to the percentage

Evaluation Method of Low-Carbon Operation to Freeway Service Zone

Because of the data collected of service is small and subjective, to improve the reliability and convenience of the model, the article use the way of AHP and TOPSIS. AHP method is used to calculate the weight of the rule-level index to target level, and the TOPSIS method is used to sort the object of evaluation.

Weight of Index Obtained by AHP Method

Using AHP method to determine the weight of different index to the target of low-carbon operation to freeway service zone, the detail procedure is as follow:

1. To build hierarchical structure of target level. The level mainly shows the goal, factors, and the relationship of different objects. The target level is represented by A, the first index level is represented by B, the second level is represented by B, and the application level is represented by E.
2. To build judged matrix. The elements of the same level are compared with each other, then, the importance of one element is obtained. The article invited 15 exports to grand the judged matrix by the way of 1–9 scale proposed by Saatty

(1980). The arithmetic average value of the marks is the element of judged matrix, then, the judged matrix A and B, which mean the first index level to the target level and the second index level to the first index level is built. The element a_{ij} in matrix A means the importance of index B_i to index B_j.

3. Hierarchical order and consistency checking. The procedure mainly calculated the order value which means relative importance of the same level index to one of previous level index, which also is the weight. The way used is standard column average method, as follows:

To normalize the elements in matrix A:

$$a_{ij} = \frac{a_{ij}}{\sum_{k=1}^{n} a_{kj}} (j = 1, 2, \cdots, n) \tag{1}$$

Then, the normalized matrix A is obtained:

$$\bar{A}' = \begin{bmatrix} a_{11} & \cdots & a_{1n} \\ \vdots & \ddots & \vdots \\ a_{n1} & \cdots & a_{nn} \end{bmatrix}$$

To calculate the sum of the elements in each column:

$$\omega_i = \sum_{j=1}^{n} a_{ij} (i = 1, 2, \ldots n) \tag{2}$$

To normalize the w_i;

$$W' = (\varpi_1, \varpi_2, \ldots, \varpi_n)^T \tag{3}$$

The vector W is the weight vector, which means the weight of first index level to the total target level. In the same way, the weight vector W_i'' can be obtained:

$$W_i'' = (\varpi_{1i}, \varpi_{2i}, \ldots, \varpi_{mi})^T \tag{4}$$

The weight of second index level to first level is

$$W'' = (W_1'', W_2'', \ldots, W_n'')^T \tag{5}$$

In the vector, n is the number of index, m is the number of the second index in first index level B_i. However, the judged matrix is built by the marks signed by the exports, there will be contrary when every two elements comparing. It is necessary to check the random consistency CR. If value of CR < 0.1, it means the matrix passes the CR checking, otherwise, the matrix is needed to adjusted. CI is the single row consistency index, and RI is the random consistency index. The formula as follows:

$$CI = \frac{\lambda_{\max} - n}{n - 1}, \tag{6}$$

$$CR = CI/RI \tag{7}$$

4. Total taxis of hierarchy. The goal is to determine the weight of the second index level to target level, the weight vector is W:

$$W = W' \cdot W'' = (W_1, W_2, \ldots, W_i, \ldots, W_n) \tag{8}$$

Therefore, the vector W is the weight coefficient by TOPSIS method.

Low-Carbon Operation Evaluation of Service Areas Based on TOPSIS

1. Evaluation matrix.
   Create an evaluation matrix consisting of m alternatives and n criteria, with the intersection of each alternative and criteria given as x_{ij}. The evaluation matrix X is:

$$X = \begin{bmatrix} x_{11} & \cdots & x_{1n} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{mn} \end{bmatrix},$$

x_{ij} represents the intersection of service area m ($i = 1, 2,\ldots, m$) and criteria j ($j = 1, 2,\ldots, n$)

2. The matrix X is then normalized to form the matrix.
   In order to remove the dimension difference, the data are standardized. X is normalized to form the matrix Y by using the normalization method.

$$y_{ij} = x_{ij} / \sqrt{\sum_{i=1}^{m} x_{ij}^2} (i = 1, 2, \ldots, m; \ j = 1, 2, \ldots, n) \tag{9}$$

3. Calculate the weighted normalized decision matrix Z.

$$z_{ij} = y_{ij} \cdot W_j$$

where ($j = 1, 2,\ldots,n$), W_j is the original weight given to the indicator.

4. Determine the best alternative Z^+ and the worst alternative Z^-. $Z^+ = (z_1^+, z_2^+, \ldots, z_n^+)$, $Z^- = (z_1^-, z_2^-, \ldots, z_n^-)$ 其中:

$$Z_j^+ = \max\left(z_{1j}, z_{2j}, \ldots, z_{mj}\right)(j = 1, 2, \ldots, n) \tag{10}$$

$$Z_j^- = \min\left(z_{1j}, z_{2j}, \ldots, z_{mj}\right)(j = 1, 2, \ldots, n) \tag{11}$$

5. Calculate the distance.
   The distance between the target alternative i and the best condition:

$$D_i^+ = \sqrt{\sum_{j=1}^{n} \left(z_{ij} - z_j^+\right)^2}(i = 1, 2, \ldots, m) \tag{12}$$

The distance between the target alternative i and the worst condition:

$$D_i^- = \sqrt{\sum_{j=1}^{n} \left(z_{ij} - z_j^-\right)^2}(i = 1, 2, \ldots, m) \tag{13}$$

6. Rank the different service area alternatives.

Calculate the service area for the optimal solution of relative proximity and ranking.

$$E_i = D_i^- / D_i^+ + D_i^- \tag{14}$$

According to relative proximity, E is more close to one and scheme is optimized when E is less than one. It is shown that evaluation level of operation low carbon is higher.

Application

Determination the Weight of Indicators Based on AHP

1. The judgment matrix of goal layer based on decision makers

$$A = \begin{bmatrix} 1 & 1/2 & 1/5 & 4 \\ 2 & 1 & 1/5 & 1/3 \\ 1/5 & 1/5 & 1 & 1/3 \\ 4 & 1/3 & 1/3 & 1 \end{bmatrix}$$

2. According to Eqs.: 1–7
   Primary index weight vector of the target layer

$$W' = (0.4790, 0.0676, 0.3141, 0.1392)^T,$$

To calculate the value of CR:

$$CR = 0.0421,$$

This value is less than 0.1, so the consistency is good, and it can be accepted.

3. The judgment matrix of goal layer based on decision makers

$$B_1 = \begin{bmatrix} 1 & 1/4 & 1/4 & 1/6 & 1/6 & 1/4 \\ 1/4 & 1 & 5 & 1/2 & 3 & 6 \\ 1/4 & 5 & 1 & 1/5 & 1/2 & 2 \\ 1/6 & 1/2 & 1/5 & 1 & 3 & 5 \\ 1/6 & 3 & 1/2 & 3 & 1 & 3 \\ 1/4 & 6 & 2 & 5 & 3 & 1 \end{bmatrix}$$

$$B_2 = \begin{bmatrix} 1 & 1/2 & 3 \\ 5 & 1 & 1/6 \\ 1 & 1/6 & 1 \end{bmatrix} B_3 = \begin{bmatrix} 1 & 5 & 1 \\ 5 & 1 & 1/6 \\ 1 & 1/6 & 1 \end{bmatrix}$$

$$B_4 = \begin{bmatrix} 1 & 5 & 1/2 & 3 \\ 5 & 1 & 2 & 7 \\ 1/2 & 2 & 1 & 3 \\ 3 & 7 & 3 & 1 \end{bmatrix}$$

The same to above, second index weight vector of the target layer is determined as W_j''. The consistency of judgment matrix CR is, respectively, 0.0089, 0.0036, 0.0026和0.0053 (CR < 0.1). Therefore, CR values are satisfied to consistency.

According to Eq. (8), the second index weight vector of the target layer is $W = (0.1099, 0.0963, 0.0837, 0.0212, 0.0357, 0.1822, 0.0209, 0.0393, 0.0074, 0.9583, 0.0263, 0.9596, 0.0258, 0.0137, 0.0611, 0.0385)^T$ defined as weight of TOPSIS evaluation.

Low-Carbon Evaluation of Three Service Areas Based on TOPSIS

1. Evaluation index of three service areas (Table 2).
2. The matrix X is normalized to form the matrix Y according to (9), weights is then determined by AHP to calculate decision matrix Z in Table 3.
3. Determine the best alternative Z^+ and the worst alternative Z^- of three service areas according to Eqs. (10) and (11).

Table 2 Evaluation index of service operation low carbon

	C_1	C_2	C_3	C_4	C_5	C_6	C_7	C_8	C_9	C_{10}	C_{11}	C_{12}	C_{13}	C_{14}	C_{15}	C_{16}
JX	0.63	3.63	0.31	0.41	3.1	0.56	1.01	3.1	0.62	0.31	0.21	0.11	1.31	3.13	0.61	1.31
XZ	0.71	2.26	0.43	0.53	2.14	0.41	2.14	2.3	0.75	0.13	0.25	0.15	2.56	2.57	0.53	2.32
LH	0.73	1.57	0.81	0.49	4.12	0.67	4.26	4.12	1.51	0.54	0.36	0.03	4.51	1.51	0.51	0.33

Table 3 Decision matrix Z

0.002	0.013	0.028	0.007	0.005	0.040	0.012	0.050	0.008	0.008	0.019	0.005	0.001	0.003	0.014	0.013
0.002	0.008	0.034	0.015	0.003	0.048	0.014	0.044	0.017	0.002	0.036	0.004	0.002	0.003	0.004	0.014
0.004	0.006	0.027	0.028	0.006	0.033	0.008	0.057	0.028	0.016	0.025	0.005	0.002	0.003	0.018	0.013

Table 4 Evaluation results of three service areas

	D^+	D^-	E
JX	0.0041	0.0015	0.27
XZ	0.0034	0.0019	0.36
LH	0.0009	0.0051	0.84

$$Z^+ = (0.004, 0.013, 0.034, 0.028, 0.006, 0.048, 0.014, 0.057, 0.028,$$
$$0.016, 0.036, 0.005, 0.002, 0.003, 0.01, 8, 0.014)$$
$$Z^- = (0.002, 0.006, 0.027, 0.007, 0.003, 0.033, 0.008, 0.044, 0.008,$$
$$0.002, 0.019, 0.004, 0.001, 0.003, 0.004, 0.013)$$

4. Calculate the distance D^+ and D^-.
   According to Eqs. (12) and (13), the distance between target and the best solution/worst solution is shown in Table 4.
5. According to Eq. (14), the results are shown in Table 4. The operation low-carbon level of LH service area is the highest, and JX operation level is lowest.

Conclusion

This paper presented the evaluation of operation in service areas of expressway by reconsidering TOPSIS and APH method used to evaluate the service level of low-carbon operation and by analyzing the interplay between service area's function and carbon emission. From two-dimensional perspective of energy conservation and emissions reduction, evaluation system of low-carbon operation in service area is constructed by basic data and evaluation model.

1. The evaluation model is scientific and reliability for the study of service area of low-carbon operation theory, and management practice provides the reference in expressway.
2. From the point of low-carbonization evaluation results, clean energy utilization and energy-saving lamps share a greater influence on the low-carbon operation of expressway service area. Service area, therefore, should as far as possible to improve the clean energy utilization rate and utilization rate of energy-saving lamps, such as the development of solar thermal, photovoltaic power generation, wind-light complementary lighting, ground source heat pump air conditioning system, and replace the green lighting equipment, achieving service daily operation and energy saving, environmental protection, and economy.

References

1. Wang, Wei. 2008. *Research on expressway rest area operation management*. China: Wuhan University of Technology.
2. Wang, Cheng. 2011. *Study on green highway service area assessment system-take highway service area of Shanxi province as an example*. Xi'an: Xi'an Chang'an University.
3. Hou, Ting-ting, Xu, Dong, and Fen HE, et al. 2013. Constructed wetland ecosystem valuation of highway service area. *Environmental Science & Technology* (12): 412–415.
4. li, Xiao-qin, and Yuan, Yin. 2012. A conceptual model of low carbon scenic spots and construction of evaluation index system. *Tourism Tribune* (03): 84–89.
5. Zhang, Chen. 2013. *Expressway service area business development and function extension*. Xi'an: Xi'an Chang'an University.
6. Chen, Qiu-hua, and Xiao-min, Zheng. 2013. Study on construction of evaluation index system of forest tourism in lowcarbon. Fujian Tribune (*The Humanities & Social Sciences Bimonthly*) 01: 38–42.
7. The "Twelve-Five" development planning of transportation. 2011. Bei Jing: The Ministry of transport.
8. Hu, Sai-yang. 2014. Construction of evaluation index of traffic system on the basis of low-carbon development. *Science and Technology Management Research*, 233–237.
9. Grazi, F., C.J.M. van den Jeroen, and Bergh. 2008. Spatial organization, transport, and climate change: Comparing instruments of spatial planning and policy. *Ecological Economics* 67 (4): 630–639.
10. Zhu, Guo-xing, and Jin-lian, Wang, et al. 2013. Research on evaluation index system of low-carbon tourism in mountain scenic areas: A case study of Huangshan scenic area. *Geographical Research*, 32.
11. Gong, Yi-long, and Da-quan, Wei. 2012. Design and application of Low-carbon road traffic evaluation model-Taking Heilongjiang Province for example. *Wuhan University Journal* 06: 799–803.
12. Duan, Yong-feng, and Hai-xia, LUO. 2014. Evaluation on resource -based cities' low- carbon economy based on dea-case study of resource—based cities in inner Mongolia autonomous regions. *Science and Technology Management Research* 01: 234–238.

An Analysis of the Taxi-Sharing Organizing and Pricing

Kai Huang, Feifei Liu, Yimian Hu and Zihan Liu

Abstract At present, the supply of taxi service is larger than the demand in many cities, particularly in mage cities. Increasing the quantity of taxi is an unsustainable method, which will aggravate the congestion and damage the environment. Therefore, in order to solve this tricky problem, we encourage passengers who have the taxi trip demand to share taxi with others. Taxi sharing has the broad prospect, but it has not been paid enough attention. Studies show that the researches of the taxi-sharing organizing (including the taxi-sharing matching and the route choice) and the pricing method are not mature and deeply. Hence, in terms of their important functions to promote the taxi sharing, this paper focuses on two aspects: the taxi-sharing organizing model and the pricing method. Firstly, we establish a based taxi-sharing organizing (the matching and route choice) model. Its objective function is to minimize the number of service vehicles and minimize the total travel distance. Furthermore, in order to overcome the limitations of the based model, we establish an improved taxi-sharing organizing (the matching and route choice) model. Its objective functions are to minimize the number of service vehicles and maximize the overlapped travel distance. This improved model can fulfill the system optimization further. Secondly, the fair fare allocation is raised. The order to pay, the number of sharing passengers, the travel distance, and the waiting time are considered, respectively. Furthermore, based on the fair fare allocation, we proposed a generalized taxi-sharing pricing method. Finally, in the viewpoint of companies/agents, drivers, and passengers, we make some suggestions to put the above two models/methods into practice and promote the taxi sharing.

K. Huang · Y. Hu · Z. Liu
Jiangsu Key Laboratory of Urban ITS, Jiangsu Province Collaborative
Innovation Center of Modern Urban Traffic Technologies,
Southeast University, Nanjing 210096, China
e-mail: h.uangkai@foxmail.com

F. Liu (✉)
Rongcheng Campus, Harbin University of Science and Technology,
Rongcheng 264300, China
e-mail: liufer-2005@163.com

Introduction

In recent years, the taxi sharing is becoming a common travel mode. The increasing number of passengers and drivers believe that sharing taxi with others is beneficial to the environment, the traffic congestion, and the participants [6, 13, 19, 23]. For the environment, sharing taxi can reduce the use of the fuel oil and the emission of the tail gas. For the society, taxi sharing can satisfy the huge trip demand. It is beneficial to decrease travelers' dissatisfaction with the public traffic, in particular when they cannot get the taxi service for a long time. For the passengers, sharing one taxi with others can save her or his travel cost, and it is a win–win cooperation. Normally, the more sharing passengers, the lesser travel cost of per participant. For drivers, they can gain more when there are more passengers taking the taxi.

Compared with the prevailing taxi-sharing practices, the researches of taxi-sharing organizing and pricing are limited. The taxi-sharing phenomenon of waving at the road side and bargaining is popular in some cities, especially in the developing countries [13]. Those immature organizing measures and pricing methods will limit its development. Therefore, studying the taxi-sharing organizing (the matching and route choice) as well as the fair pricing is necessary to promote its development and build a health sharing market.

The two procedures of the taxi-sharing matching and the route choice are related and interactive. When the taxi-sharing organizers (the taxi companies or the agents) arrange those waiting passengers to share taxi, they have to consider the sharing route. If only considering the matching procedure with the nearest destinations, the selected route may be not the best one (the shortest travel route). Therefore, we combine the two procedures and try to establish a generalized organizing model.

Studying the pricing method needs to fulfill the fair fare allocation in advance [20, 22–24]. The taxi-sharing enthusiasm of passengers will be influenced by the sharing charge. Due to the difference of service order, the travel distance, the numbers of sharing passengers, and the waiting time, the cost-sharing ratio should be changeable. Then, based on those factors, establishing a generalized pricing method is necessary.

The first aim of this paper is to make a clear review of the researches of the taxi sharing, which includes two main aspects: the taxi-sharing organizing and the pricing method. The second aim is to establish the taxi-sharing organizing (the taxi-sharing matching and the route choice) model and analyze those potential limitations within this model. If it does, we will establish an improved model which will be more persuasive. The third aim is to study the fair pricing including the fair allocation and the practical pricing method. From the participants' point of view (the company/agent, the passenger, and the driver), the last aim is to make some suggestions, which are based on the above model/method.

The remainder of the paper is organized as follows: Section "Literature Review" introduces the main researches related to the taxi-sharing organizing (the matching and route choice) as well as the pricing, Section "The Taxi-sharing Matching and Route Choice Model" establishes the taxi-sharing models: a based model and an improved model, Section "The Taxi-sharing Pricing Method" analyzes the equity issue of taxi sharing and raises a pricing method, Section "Suggestions" makes some suggestions to promote the taxi sharing, and Section "Conclusion and Future Research" concludes.

Literature Review

The related research of the taxi-sharing organizing (the taxi-sharing matching and route choice) mainly focuses on the modeling of the routing [9, 10]. The objective functions are to maximize the income of the driver, maximize the society profit, minimize the travel cost, minimize the travel time, and minimize the operation cost.

Wu et al. [32] proposed a taxi-sharing route choice model, whose objective functions are to maximize the society profit and minimize the travel cost of passengers. And they used the Hybrid Particle Swarm algorithm to solve this model. The optimization result coming from the simulation shows that utilizing the Hybrid Particle Swarm algorithm is an effective and accurate method. In 2011, Zhou et al. established a taxi-sharing route choice model, whose objective functions are to minimize the travel cost of passengers and ensure the minimum requirement of drivers' income. They utilized the Heuristic algorithm to solve this mixed integer programming. The result shows that this Heuristic algorithm is a time-consuming but accurate method. Lin et al. [18] raised an optimization model of vehicle route program for ride-sharing taxi. The objective functions are to minimize the operation cost and maximize the satisfaction of passengers. The simulated annealing was used to solve this program, and the result shows that the total travel distance and the number of serviced taxi are lower. In 2013, Santos and Xavier used the Contraction Hierarchy algorithm to solve this dynamic dial-a-ride program, whose objective functions are to maximize the number of sharing passengers and minimize the travel charge of passengers. It is found that this algorithm can give a more accurate calculation. Hosni et al. [16] used the Lagrangian relaxation to solve the taxi-sharing routing problem, whose objective functions are to maximize the average income of passengers and minimize the travel charge of passengers. And it is found that the Lagrangian approach has the higher efficiency.

The related researches of the taxi-sharing pricing mainly focus on two aspects: the equity issue of the fare allocation and the pricing method. For the equity issue of the fare allocation, many researchers give lots of unique ideas. Jorge et al. [17] suggested assigning the sharing fare averagely. This allocation method is based on the number of passengers. In 2015, Santos and Xavier suggested taking the travel distance as the key factor when calculating the fare allocation. This kind of allocation method is fairer to the sharing passengers.

After deciding the allocation method, the generalized pricing method is easy to establish. Normally, the taxi-sharing pricing method is similar to the taxi pricing. There is a positive relationship between the sharing charge and the travel distance [11, 25]. Also, in terms of the sharing, there is a negative relationship between the personal charge and the number of passengers. In addition, Zikeš [35] proposed a new allocation method, which takes the trip urgency into consideration. When the waiting passenger wants to get sharing service in advance, she or he has to pay more. Coincidentally, the popular taxi calling service providers such as DIDI or Uber have the similar services: the VIP service with higher charge [14].

Seeing some limitations of the taxi-sharing matching and route choice as well as the pricing method, it is necessary to find out some optimization models/methods. The main contributions of this paper are the followings. (a) This paper establishes a based and an improved taxi-sharing organizing (including the two procedures of the taxi-sharing matching and the route choice) model. The two models are suitable for two and more passengers to share one taxi, which will be more complicated and difficult than the sharing model for only two passengers. (b) We take the order to pay, the number of sharing passengers, the travel distance, and the waiting time into consideration when fulfilling the fair fare allocation. Furthermore, a practical pricing method is proposed.

The Taxi-sharing Matching and Route Choice Model

The taxi-sharing matching is to find the sharing passengers and establish a matching model. The model is based on the objective function of minimizing the number of service vehicles. The taxi-sharing route choice is to find the sharing route and establish a routing model. The model is based on the objective functions of minimizing passengers' travel cost, minimizing total travel distance, or minimizing the travel distance. As discussed in Section "Introduction", the two procedures (the matching and the routing) are complementary to each other.

In this paper, we combine the two procedures and establish the generalized models. The models are suitable for 2–4 or more sharing passengers serviced by one taxi. The following is the introduction of the new models.

The Based Taxi-sharing Matching and Route Choice Model

In this model, the objective functions include two parts: minimizing the number of serviced taxi and minimizing the travel distance of each sharing passenger. Suppose that, there are n passengers waiting for sharing taxi in the sharing station (it may be the railway station, the bus station, or other waiting place), and there are V vehicles waiting for providing the taxi-sharing service. The M is a large positive value. The D_i is the total sharing distance of the ith passenger. The D_{ji} is the distance between

the destinations of the jth passenger and the ith passenger (if the $j = o$, the D_{oi} denotes the distance from the origin to the ith passenger' destination). The δ_i^v is a binary decision variable that equals 1 if the passenger i is serviced with the vth vehicle, and 0 otherwise. The σ_{ji}^v is a binary decision variable that equals 1 if vehicle v serves the jth passenger in advance, then the ith passenger (travel from the destination of j to i), and 0 otherwise. The T_i is the travel time of the ith passenger based on the prediction of the travel distance. The $\overline{T_i}$ is the average travel time of the ith passenger based on the historical data. The following is the model:

$$\text{Min}\left\{ M \sum_{v=1}^{V} \left[\left(4 - \sum_{i=1}^{n} \delta_i^v \right) \sum_{i=1}^{n} \delta_i^v \right] + \sum_{i=1}^{n} D_i \right\} \tag{1}$$

$$D_i = \left(1 - \sum_{j=1}^{n} \sigma_{ji}^v \right) D_{oi}$$

$$+ \sum_{j=1}^{n} \sigma_{ji}^v \left\{ \begin{array}{l} \left(1 - \sum_{l=1}^{n} \sigma_{lj}^v \sum_{j=1}^{n} \sigma_{ji}^v \right) \left(D_{oj} + D_{ji} \right) \\[2em] + \sum_{l=1}^{n} \sigma_{lj}^v \sum_{j=1}^{n} \sigma_{ji}^v \left[\begin{array}{l} \left(1 - \sum_{k=1}^{n} \sigma_{kl}^v \sum_{l=1}^{n} \sigma_{lj}^v \sum_{j=1}^{n} \sigma_{ji}^v \right) \left(D_{ol} + D_{lj} + D_{ji} \right) \\[1.5em] + \sum_{k=1}^{n} \sigma_{kl}^v \sum_{l=1}^{n} \sigma_{lj}^v \sum_{j=1}^{n} \sigma_{ji}^v \left(D_{ok} + D_{kl} + D_{lj} + D_{ji} \right) \end{array} \right] \end{array} \right.$$

$$\tag{2}$$

subject to:

$$\sum_{i=1}^{n} \delta_i^v \leq 4 \quad v = 1, 2 \ldots, V \tag{3}$$

$$\sum_{v=1}^{V} \sum_{j=1}^{n} \sigma_{ji}^v \leq 1 \quad i = 1, 2 \ldots, n \tag{4}$$

$$\sum_{v=1}^{V} \sum_{j=1}^{n} \sigma_{ij}^v \leq 1 \quad i = 1, 2 \ldots, n \tag{5}$$

$$\sum_{l=1}^{n} \sigma_{li}^v + \sum_{j=1}^{n} \sigma_{ij}^v \leq M \delta_i^v \quad i = 1, 2 \ldots, n \ \ v = 1, 2 \ldots, V \tag{6}$$

$$\sum_{v=1}^{V} \delta_i^v = 1, \quad i = 1, 2 \ldots, n \tag{7}$$

$$\sigma_{ji}^v \sigma_{ij}^v = 0 \quad i,j = 1,2\ldots,n \ \ v = 1,2\ldots,V \tag{8}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}^v \geq \sum_{i=1}^{n}\delta_i^v - 1, \quad v = 1,2\ldots,V \tag{9}$$

$$\frac{T_i}{\overline{T_i}} < \tau \quad i = 1,2\ldots,n \tag{10}$$

The objective function (1) is to minimize the number of service vehicles and allocate passengers with the shortest travel distance to one vehicle, where the parameter M determines the weight of the two terms (minimizing the number of service vehicles should be fulfilled in advance). The (2) is a calculation model of passengers' travel distance. All the constraints are to ensure the accuracy of the model.

The Improved Taxi-sharing Matching and Route Choice Model

However, this based model has some limitations when choosing the sharing route. This based model is helpful to find out the shortest travel distance for each passenger. While, it ignores that how to minimize the average travel cost. Taking two passengers to share one taxi, for example, we suppose that the travel distance of the first passenger is too short to calculate, and the sharing charge of the second passenger will be close to the non-sharing travel cost. Therefore, we proposed an improved taxi-sharing matching and the route choice model. Its objective functions are to minimize the number of serviced taxi and maximize the overlapped travel distance of sharing passengers. The following is the improved model:

$$\text{Min}\left\{ M\sum_{v=1}^{V}\left[\left(4 - \sum_{i=1}^{n}\delta_i^v\right)\sum_{i=1}^{n}\delta_i^v\right] + \sum_{v=1}^{V}\sum_{i=1}^{n}\left(\sum_{l=1}^{n}D_{li}\sigma_{li}^v + \sum_{j=1}^{n}D_{ij}\sigma_{ij}^v\right)\right\} \tag{11}$$

subject to:

$$\sum_{i=1}^{n}\delta_i^v \leq 4 \quad v = 1,2\ldots,V \tag{12}$$

$$\sum_{v=1}^{V}\sum_{j=1}^{n}\sigma_{ji}^v \leq 1 \quad i = 1,2\ldots,n \tag{13}$$

$$\sum_{v=1}^{V}\sum_{j=1}^{n}\sigma_{ij}^{v} \leq 1 \quad i = 1, 2. . ., n \tag{14}$$

$$\sum_{l=1}^{n}\sigma_{li}^{v} + \sum_{j=1}^{n}\sigma_{ij}^{v} \leq M\delta_{i}^{v} \quad i = 1, 2. . ., n \quad v = 1, 2. . ., V \tag{15}$$

$$\sum_{v=1}^{V}\delta_{i}^{v} = 1, \quad i = 1, 2. . ., n \tag{16}$$

$$\sigma_{ji}^{v}\sigma_{ij}^{v} = 0 \quad i, j = 1, 2. . ., n \quad v = 1, 2. . ., V \tag{17}$$

$$\sum_{i=1}^{n}\sum_{j=1}^{n}\sigma_{ij}^{v} \geq \sum_{i=1}^{n}\delta_{i}^{v} - 1, \quad v = 1, 2. . ., V \tag{18}$$

$$\frac{T_i}{\overline{\overline{T_i}}} < \tau \quad i = 1, 2. . ., n \tag{19}$$

Compared with the model (1), the second objective function of model (11) is to minimize the distance of sharing passengers' destinations, which means maximizing the overlapped travel distance and minimizing the average travel cost.

The Taxi-sharing Pricing Method

One of the main aims is to save the travel cost when the passengers wait for sharing taxi. Hence, the allocation method of sharing charge will be important to ensure the vitality of taxi sharing. The participation will be influenced by the fair fare allocation. To raise a generalized pricing method, the first step is to consider the equity issue. The allocation method of the taxi-sharing charge should be different with the change of some factors, such as the increase of the travel distance due to the detour, the increase of the travel time due to the more getting on or off, and the lower comfort due to the more passengers in one taxi. The second step is to establish the generalized pricing model, which includes the number of passengers and the travel distance.

The Equity Issue of the Fare Allocation

To ensure taxi-sharing equity, there are four scenarios considered in the literature [11, 25]. The four scenarios include the payment in turn, the payment based on the

number of sharing passengers, the payment based on the travel distance, and the payment considering the waiting time.

The Payment in Turn

In the early period of the taxi sharing, it often occurs in the acquaintances. After sharing taxi, one passenger will pay the bill voluntarily, and others in turn. This kind of pricing method has some benefits: (a) the efficiency of payment is high, and it is unnecessary to calculate per charge; (b) it is easy to organize sharing taxi, because that those familiars will wait for searchers take the pricing strategy into the taxi at a fixed location. However, the fare allocation has some disadvantages: (a) it is clearly not fair, as the charge for the late-served passengers is obviously lower; (b) it is unfair for those passengers to pay the same bill whose travel distance are shorter than others.

The Payment Based on the Number of Passengers

To overcome the limitation of the above fare allocation method, the allocation method based on the number of sharing passengers is proposed. The sharing fare of per sharing passenger is the same. The following is the fare allocation method:

$$\frac{F_i}{\sum_{j=1}^{N} F_j} = \frac{1}{N} \quad i,j = 1,2,\ldots,N \quad N = 2,3,4 \tag{20}$$

where F_i is the sharing charge of the ith passenger, N is the number of passengers who share one taxi.

The fare allocation method based on the number of passengers has some benefits:

(a) it is easy to calculate the charge of per passenger, and it is the same for everyone;
(b) it overcomes the first limitation discussed in the payment in turn, and the fare of the early-served is equal to the late-served. However, the second limitation discussed in the payment in turn still be not avoided. Obviously, it is unfair for the passenger whose travel distance is shorter.

The Payment Based on the Travel Distance

In terms of limitation caused by the different travel distances, the fare allocation method based on the travel distance is raised. The percentage of one passenger

should bear is proportional to his/her travel distance. The following is the fare allocation method:

$$\frac{F_i}{\sum_{j=1}^{N} F_j} = \frac{\bar{D}_i}{\sum_{j=1}^{N} \bar{D}_j} \quad i,j = 1, 2, \ldots, N \quad N = 2, 3, 4 \tag{21}$$

where F_i is the sharing charge of the ith passenger. $\bar{D}_i$ is the travel distance of the ith passenger, which is based on the historical data.

This allocation method overcomes most of the limitations discussed in Sections "The Payment in Turn" and "The Payment Based on the Number of Passengers". However, it will be more time-consuming and high-cost to collect historical data. In the future, with the popularity of electronic map, it will be easy to get the data and ensure its accuracy.

The Payment Considering the Waiting Time

One non-ignorable problem is that the waiting time may be longer for those passengers who arrive first than those latecomers. If her or his waiting time is too longer, the passenger should be given a discount. We proposed an improved allocation method based on the method in Section "The Payment Based on the Travel Distance" (the payment based on the travel distance). The following is the fare allocation method:

$$\frac{F_i}{\sum_{j=1}^{N} F_j} = \frac{\lambda \bar{D}_i}{\sum_{j=1}^{N} \bar{D}_j} \quad i,j = 1, 2, \ldots, N \quad N = 2, 3, 4 \tag{22}$$

where λ is the discount factor, and the recommended value belongs to [0.8, 1.0]. The value will be smaller if the waiting time is too long.

The Generalized Pricing Method of the Taxi Sharing

There are a lot of related researches of the taxi-sharing pricing method, which is influenced by many factors. In 2012, Hong raised a taxi-sharing pricing method based on the supply and demand of taxi, the citizens' income, the operation cost, and the competition of several trip modes. Zhang et al. [33] took the benefits of the passenger, the driver, the taxi company, and the society into consideration when raising the pricing method. In addition, some researchers take the pricing strategy into a specific application. Some related pricing devices/patents have been developed which can bring lots of benefits to users [4, 5, 7].

We find that the above taxi-sharing pricing methods are miscellaneous, and most of them are suitable for the limited cases. Studies show that there are four main

Fig. 1 Four main
taxi-sharing cases
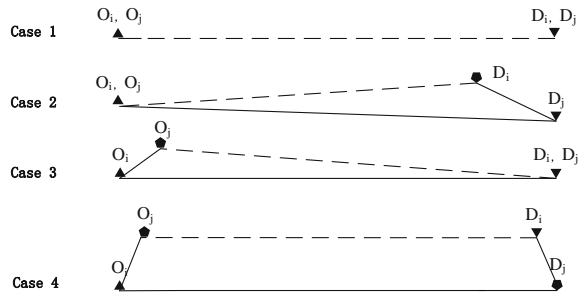
taxi-sharing cases: (a) passengers have the same origin and the same destination, (b) passengers have the same origin and the different destinations, (c) passengers have the different origins and the same destination, (d) passengers have the different origins and the different destinations. Two passengers sharing one taxi is the basic model, and the pricing method of three or four passengers is similar. The following is the diagram (Fig. 1):

In terms of the above four main cases, we propose a generalized pricing method, which includes two parts: (a) we take the sharing pricing as the first part (the fare allocation method in Section "The Payment Considering the Waiting Time" is recommended), (b) we take a fixed operation cost as the second part. The following is the generalized pricing method:

$$F_i = \frac{\lambda \bar{D}_i}{\sum_{j=1}^{N} \bar{D}_j} \bar{F} + f \quad i,j = 1,2,\ldots,N \quad N = 2,3,4 \tag{23}$$

where $\bar{F}$ is the total fare for the taxi sharing, which is coming from the historical data. f is the fixed operation cost of the taxi sharing. Note that adding f is to benefit the driver during the taxi sharing.

Suggestions

The taxi sharing needs the cooperation of participants, which includes three parts of the company/agent, the passenger, and the driver. The company/agent is the decision maker, who will match those waiting passengers, choose the sharing route, and select the pricing method. The driver and the passenger are the executors, who will put the taxi-sharing matching and route choice model as well as the pricing method into practice. This section will give a discussion that what the company/agent, the passenger, and the driver will do to promote the taxi sharing.

The Taxi-sharing Organizing Assisted by a Company/Agent

A taxi-sharing company/agent can promote the taxi-sharing matching services [2]. It can help the taxi dispatching center or taxi-sharing platform to communicate with drivers and passengers [8, 12, 29, 30].

For all sorts of taxi calling services, the drivers and passengers usually complain about the bad matching allocations from the center/platform. Hence, such a company/agent can help to get better matching plans: allocate the available taxis to the waiting passengers who are closer. It can significantly reduce the waiting time of passengers and the empty cruising time of taxi. Such a better plan can be obtained based on the analysis of historical data, as well as getting more feedbacks from the users of the sharing services. Also, it is important for such a company/agent to play as a third party between the users and the system/platform, in order to help the two sides on the communications and also to some extent custody their behaviors.

The Call-a-Ride Used by the Driver and the Passenger

With the development of Information Technology (IT), the call-a-ride will be the main taxi-sharing model [1, 14, 26, 31]. Recently, the taxi-sharing APP is popular, which has the functions of the matching, the route choice, and the pricing. Passengers can send sharing information with their mobile phones. Then, the dispatching center (the taxi company/agent) will dispatch taxi, choose sharing route, and calculate sharing price.

The study shows that the young are the main target population. It may be that the young are more willing to use the smartphone. Aoun et al. [3] found that the young students prefer utilizing taxi-sharing APP to call a taxi rather waiting at the roadside. Usually, they use APPs to share taxi with strange students, and the success rate is higher than waving down a taxi at the roadside. Therefore, we encourage the driver and passenger to use call-a-ride (the taxi-sharing APP), which will enhance the sharing efficiency.

Conclusion and Future Research

Taxi sharing can bring many benefits to the environment, economy, traffic congestion, and society. In 2013, Liu et al. found that over 63% of passengers wants to share taxi with others in Shanghai, which is the biggest city (the population and the economy) in China. While, it is not optimistic in less-developed cities. In African, it is a common phenomenon that the motorbike taxi resists the development of taxi sharing [13]. Therefore, this paper establishes the organizing (the taxi-sharing

matching and route choice) model and the pricing method to promote its development. Afterward, some suggestions are made to encourage the company/agent to organize the taxi sharing, and the passenger and the driver to use the call-a-ride.

In this paper, we establish two taxi-sharing matching and route choice models. The based model is to minimize the number of serviced taxi and minimize the travel distance of each sharing passenger. While, it cannot minimize the average travel cost. Then, an improved model is raised, which is to minimize the number of serviced taxi and maximize the overlapped travel distance of sharing passengers. Besides, we analyze the equity issue of fare allocation. The order to pay, the number of sharing passengers, the travel distance, and the waiting time are considered. Furthermore, a generalized pricing model is proposed, which is based on the fair fare allocation. Finally, we make the company/agent, the driver, and the passenger two suggestions: The company/agent organizes the taxi sharing as well as the driver and the passenger use the call-a-ride (the taxi-sharing APP).

This paper provides an initial study of the taxi sharing. However, there are some issues still remaining unsolved well. (a) For the taxi-sharing organizing model, minimizing the travel time is not paid enough attention. The model may be changed a lot if considering the travel time. (b) For the pricing method, the value of time (VOT) of passengers may be a meaningful factor influencing the fare allocation and pricing. For these people who have high VOT, they may be willing to pay more for better services. In contrast, the lower price will attract some passengers whose VOT is low. (c) For the taxi-sharing APP, some invalid information will result in the failure of taxi sharing. The reason is that most taxi-sharing APPs are free when users send sharing taxi requirements. If they find better travel modes, passengers would give up the taxi-sharing service which is booked just now.

References

1. Anderson, D. 2014. Not just a taxi?—For-profit ridesharing, driver strategies, and VMT. *Transportation* 41 (5): 1099–1117.
2. Aarhauga, J., and K. Skolleruda. 2014. Taxi: Different solutions in different segments. *Transportation Research Procedia* 1: 276–283.
3. Aoun, A., M. Abou-Zeid, I. Kaysi, and C. Myntti. 2013. Reducing parking demand and traffic congestion at the American University of Beirut. *Transport Policy* 25 (1): 52–60.
4. Cai, D. 2011. *A system of taxi pricing*. CN 202058211U.
5. Chen, L. 2013. *Pricing system of taxi sharing and order method*. CN 103150762A.
6. Chen, J., Z. Liu, S. Zhu, and W. Wang. 2015. Design of limited-stop bus service with capacity constraint and stochastic travel time. *Transportation Research Part E* 83: 1–15.

 7. Cui, J. 2011. *Pricing system of taxi sharing*. CN 202257689U.
 8. Dembele, J., and C. Cambier. 2012. An Agent-particle model for taxis-based aggregation; Emergence and detection of structures. *Procedia Computer Science* 9 (9): 1484–1493.
 9. Desaulniers, Guy, Jacques Desrosiers, Andreas Erdmann, Marius M. Solomon, and François Soumis. 2002. VRP with pickup and delivery. *The Vehicle Routing Problem* 9: 225–242.
10. Desrosiers, Jacques, Yvan Dumas, Marius M. Solomon, and François Soumis. 1995. Time constrained routing and scheduling. *Handbooks in Operations Research and Management Science* 8: 35–139.
11. Fagin, R., and J.H. Williams. 1983. A fair carpool scheduling algorithm. *Journal of Research & Development* 27 (2): 133–139.
12. Grau, J., and M. Romeu. 2015. Agent based modelling for simulating taxi services. *Procedia Computer Science* 52: 902–907.
13. Guézéré, A. 2014. The reconstruction of shared taxis as rural transport due to the competition of motor bike taxis in Togo secondary cities. *Case Studies on Transport Policy* 3: 253–263.
14. He, F., and Z. Shen. 2015. Modeling taxi services with smartphone-based e-hailing applications. *Transportation Research Part C* 58: 93–106.
15. Hong, Qilin. (2012). Research on taxi pricing based on ride-sharing. Ph.D. Thesis at *Harbin Institute of Technology*.
16. Hosni, H., J. Naoum-Sawaya, and H. Artail. 2014. The shared-taxi problem: Formulation and solution methods. *Transportation Research Part B* 70: 303–318.
17. Jorge, D., G. Molnar, and G. Correia. 2015. Trip pricing of one-way station-based carsharing networks with zone and time of day price variations. *Transportation Research Part B* 81: 461–482.
18. Lin, Y., W. Li, F. Qiu, and H. Xu. 2012. Research on optimization of vehicle routing problem for ride-sharing taxi. *Procedia—Social and Behavioral Sciences* 43: 494–502.
19. Liu, S., Z. Zhou, and C. Zhang. 2013. Research on optimization of "Taxi-pooling" in Hefei City. *Journal of Chongqing Technology and Business University.* 11 (30): 70–75.
20. Liu, Z., and Q. Meng. 2014. Bus-based park-and-ride system: A stochastic model on multimodal network with congestion pricing schemes. *International Journal of Systems Science* 45 (5): 994–1006.
21. Liu, Z., Q. Meng, and S. Wang. 2013b. Speed-based toll design for cordon-based congestion pricing scheme. *Transportation Research Part C* 31: 83–98.
22. Liu, Z., Q. Meng, and S. Wang. 2014. Variational inequality model for cordon-based congestion pricing under side constrained stochastic user equilibrium conditions. *Transportmetrica A* 10 (8): 693–704.
23. Liu, Z., S. Wang, and Q. Meng. 2014. Optimal joint distance and time toll for cordon-based congestion pricing. *Transportation Research Part B* 69: 81–97.
24. Liu, Z., S. Wang, and Q. Meng. 2014. Toll pricing framework under logit-based stochastic user equilibrium constraints. *Journal of Advanced Transportation* 48: 1121–1137.
25. Naor, M. 2005. On fairness in the carpool problem. *Journal of Algorithms* 55 (1): 93–98.
26. Rayle, L., D. Dai, N. Chan, and S. Shaheen. 2016. Just a better taxi?—A survey-based comparison of taxis, transit, and ride sourcing services in San Francisco. *Transport Policy* 45: 168–178.
27. Santos, D., & E. Xavier. 2013. Dynamic taxi and ridesharing: A framework and heuristics for the optimization problem. In *Proceedings of the 23rd International Joint Conference on Artificial Intelligence IJCAI 2013*, 2885–2891.
28. Santos, D., and E. Xavier. 2015. Taxi and ride sharing: A dynamic dial-a-ride problem with money as an incentive. *Expert Systems with Applications* 42 (19): 6728–6737.
29. Seow, K., and D. Lee. 2007. Towards an automated multiagent taxi-dispatch system. In *IEEE Conference on Automation Science and Engineering,* 31–36.
30. Seow, K., and D. Lee. 2010. Performance of multiagent taxi dispatch on extended-runtime taxi availability: A simulation study. *IEEE Transactions on Intelligent Transportation Systems* 11 (1): 231–236.

31. Song, F., R. Li, and H. Zhou. 2015. Feasibility and issues for establishing network-based carpooling scheme. *Pervasive & Mobile Computing* 24: 4–15.
32. Wu, F., Z. Li, and C. Xu. 2009. Study of the mode of combined-taxi optimal scheduling. *Journal of Lanzhou Jiaotong University* 88 (1): 104–107.
33. Zhang, W., R. He, and Q. Xiao. 2015. Research on taxi carpooling pricing multi-objective optimization. *Journal of Wuhan University of Technology (Transportation Science Engineering)* 39 (6): 1105–1109.
34. Zhou, H., B. Zhong, X. Peng, and X. Xia. 2011. The route choice and rate optimization model of taxi-pooling. *Journal of Changsha University of Science and Technology (Natural Science)* 8 (4): 20–24.
35. Zikeš, J. 2012. Auction-based taxi allocation with dynamic pricing. Bachelor Thesis of Czech Technical University in Prague.

Analyzing the Influences of Driver Distractions Based on Driver's Subjective Cognition

Hui Zhang, Da-Lin Qian, Chun-Fu Shao, Zhen-Wei Qian and Xue-Yu Mi

Abstract The study is intended to investigate the prevalence and severity of 13 types of driver distractions. Two types of survey, an anonymous online questionnaire and a field distribution and recycling questionnaire, have been performed to collect the data. Four hundred and four respondents contributed to the survey. Independent sample T-test and one-way ANOVA method are used to analyze the influences of driver personality traits (i.e., gender, professional, age, and driving experience) on the prevalence of driver distractions. In addition, clustering analysis method is used to the classification of driver distractions. Results indicated that the most frequent distraction was "Listening to music," and the most dangerous distraction was "Writing text messages." Three personality traits (professional, age, and driving experience) were found to have a significant impact on the prevalence of driver distractions. According to the influence factors, driver distractions could be divided into three categories: low, middle, and high. The multiple resource theory was used to explain the rationality of categories in the view of cognition. The results could contribute to the drivers' subjective perceptions of each driver distraction and offer the basis for the further research on driver distraction and road safety.

Keywords Driver distractions · Personality traits · Distraction classification · Multiple resource theory

Introduction

Driver distraction is now recognized as a contributing factor in at least a quarter of motor vehicle crashes [1]. This proportion, however, could increase, in line with the increased use of portable technologies such as smartphones and satellite navigation

H. Zhang · D.-L. Qian (✉) · C.-F. Shao · Z.-W. Qian · X.-Y. Mi
Key Laboratory for Urban Transportation Complex Systems Theory
and Technology, Beijing Jiaotong University, Beijing 100044, China
e-mail: dlqian@bjtu.edu.cn

systems. Therefore, it is extremely important to have a clear understanding of types of distractions that drivers currently engage in, their prevalence, and the types of drivers who are more likely to be distracted.

Although the research investigating the impact of distractions on driving performance is well advanced [2–6], the research measuring exposure to driver distractions is still in its infancy [7]. At present, there are four broad approaches to investigate this issue: self-report survey [8–11], cross-sectional survey [12, 13], roadside observation [14–16], and naturalistic observation [17–20]. Despite some scholars have used these methods for investigating drivers exposure to driver distraction, however, most of those that do exist have solely concentrated on the prevalence of one type of distracter (the mobile phone) [12, 13, 21, 22]. However, accidents caused by distracted are often the outcome of combined action of several driver distractions; therefore, research on single driver distraction cannot fully describe the reason for these accidents, which led to the prevention from the root. In this paper, according to the characteristics of the Chinese driver distraction, several types of driver distractions were summarized, and self-report survey was adopted to investigate the prevalence and severity of these driver distractions; the quantitative relationship between driver distractions and personality traits was analyzed. Finally, the clustering analysis method was used to obtain the classification of driver distractions.

Methods

A self-report survey was carried out for this study. The instrument collected anonymous self-report data via the Internet and the field distribution and recycling.

Questionnaire Design

Drivers perform many kinds of driver distractions while driving, from operated the vehicle-mounted device or mobile terminals (calls, navigation, etc.) to the driver behavior (eat food, drink, smoking, etc.). According to the crash accident data, driver distractions can be divided into 13 types by AAA [19]. Driver distractions can also be divided into 10 types based on 100-car data by NHTSA [20]. According to these two studies, combined with China's actual situation, this research comes up with 13 types of driver distractions, as shown in Table 1.

There were five sections in the survey: (i) demographics (age, gender, driving years, accident history (frequency in the previous 3 years) and profession, (ii) the frequency and severity of driver distractions (frequency: a Likert scale '1'—never to '5'—always, severity: a Likert scale '1'—not distracting to '5'—very distracting) for each item, (iii) the impact on the specific driver operation of engagement in distracting behaviors, measured by four indicators ("direction control ability,"

Table 1 Thirteen types of driver distractions

No.	Driver distractions	No.	Driver distractions
1	Answer the phone	8	Eating or drinking
2	Edit text	9	Use the navigation
3	Listen to music	10	Daydreaming
4	Chat	11	Still garbage
5	Listen to the radio	12	Moving object in vehicle
6	Operating equipment	13	Outside person, object or event
7	Smoking		

Table 2 Driver personal characteristics

Category	Gender				Professional		Age			Driving experience		
Group	Male	Female	Y	N	<25	25–40	40–50	>50	<5	5–10	10–20	>20
Number	248	156	101	303	61	202	83	58	146	157	79	22

"brake response ability," "road sign recognition ability," and "traffic forecasting ability"; each indicator used a Likert scale '1'—not distracting to '5'—very distracting).

Procedure

The questionnaire was divided into two types: Ordinary drivers and professional drivers. The ordinary drivers questionnaire survey was promoted on a professional survey Web site which is "http://www.sojump.com/." The ordinary drivers questionnaire was for private car driver. The survey was carried out between July 1, 2015, and Aug 1, 2015, then filtered for recycling questionnaire, and finally got 303 valid questionnaires. The professional drivers questionnaire was delivered to dangerous goods driver at the dangerous goods driver safety education training meeting in Shenyang in September 2015, and the survey was completed by means of field distribution and recycling. After filtering for recycling questionnaire, 101 valid questionnaires were obtained finally.

Respondents

Data were collected from 404 respondents during a month collection period, as illustrated in Table 2. In all respondents, 25% of the respondents were professional driver, 61% of the respondents were male, the average age was thirty-nine

(SD = 9.3), and the average years of driving experience was nine years (SD = 12.1).

Reliability and Validity Analysis

In order to ensure the accuracy of the survey data, the questionnaire reliability and validity analysis was necessary. In this paper, Cronbach's α was used to evaluate the questionnaire reliability. The results are shown in Table 3. The results showed that the questionnaire reliability was acceptable.

This article adopted the method of factor analysis to test the validity of the questionnaire. Firstly, KMO and Bartlett's spherical inspection was used to determine whether factor analysis was appropriate. In the KMO test, KMO value (between 0 and 1) is closer to 1, the more suitable for factor analysis. In Bartlett's spherical test, it is suitable for factor analysis when the value of Sig is lesser than 0.01. KMO and Bartlett's spherical inspection was taken for 13 driver distractions. The results are shown in Table 4.

In this paper, the value of KMO was between 0.84 and 0.88, and the Sig value of the Bartlett's spherical inspection is far less than 0.01, which is suitable for factor analysis. Factor analysis has been carried out for 13 driver distractions, and the results showed that the factor load was greater than 0.4, all through the validity test.

Results and Discussion

The Frequency and Severity of Driver Distractions

The frequency distribution of each driver distraction is presented in Fig. 1. The frequency was measured by the proportion of "always" and "often." Respondents considered that the three most frequent driver distractions were "3. Listen to music" (55.7%), "5. Listen to the radio" (50.4%), and "4. Chat" (43.5%).

The severity distribution of each driver distraction is shown in Fig. 2. The severity was measured by the proportion of "very" and "big" distracting. Respondents considered that the three most severe driver distractions were "2. Edit text" (83.6%), "12. Moving object in vehicle" (71.7%), and "13. Outside person, object or event" (68.4%).

Table 3 Reliability analysis results of 13 driver distractions

Driver distractions	1	2	3	4	5	6	7	8	9	10	11	12	13
Cronbach's α	0.76	0.78	0.71	0.76	0.67	0.77	0.81	0.80	0.77	0.81	0.85	0.86	0.83

Table 4 KMO and Blartlett test results of 13 driver distractions

Driver distractions	1	2	3	4	5	6	7	8	9	10	11	12	13
KMO	0.85	0.84	0.85	0.84	0.85	0.85	0.86	0.86	0.84	0.86	0.87	0.88	0.88
Blartlett sig. value	0	0	0	0	0	0	0	0	0	0	0	0	0

Fig. 1 Frequency distribution of each driver distraction
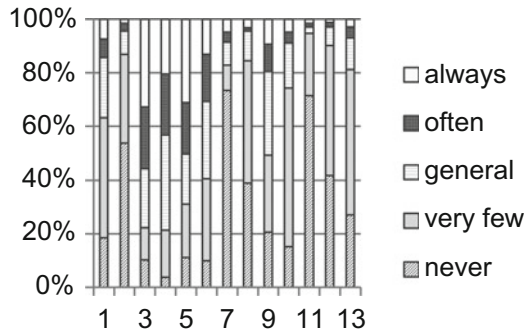
Fig. 2 Severity distribution of each driver distraction
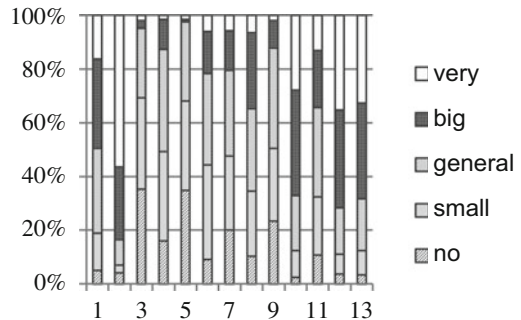
The present study investigated the frequency and severity of drivers in China who were engaged in different types of distractions while driving. The most frequent observed distractions in the present study were "Listen to music," "Listen to the radio," and "Chat." Despite of the different data collection methods, the present study also had some similarities with the previous studies on the types of distractions observed. Stutts et al. [17] installed video cameras in 70 cars in America to observe what distracted the car drivers, and they found that 92% of drivers were using voice or music equipment while driving. They found that the most common distraction was conversing with passengers, which is similar to "Chat" in the present study. Sullman [1] using six urban centers in England, talking to passengers was the most common driver distraction, followed by smoking and mobile phone use. Prat [16] investigated the prevalence of driver engagement in secondary tasks

in Spain, as the same to Sullman's results, the most common being: conversing with a passenger, smoking and talking on a handheld mobile phone. In many studies, the frequency of mobile phone use was more higher, but was not so much in this study, where the reason may be related to the use of mobile phones while driving has been banned in China.

The most severity observed distractions in the present study were "Edit text," "Moving object in vehicle," and "Outside person, object or event." "Edit text" was considered very dangerous in many researches. An Australian survey reports from RAC indicated that the top three dangerous driver distraction were (i) reading or sending text messages, (ii) attending to children, and (iii) reading maps [10]. One further survey, sampling drivers from the Australian state of Victoria, provided additional data. Excluding driving under the influence of alcohol, the respondents rated sending text messages, reading text messages, and dialing a mobile telephone as the three most dangerous driving activities [9]. Lansdown [11] used an anonymous online questionnaire to study the severity and frequency of engagement in distracting driver behaviors. Results indicated that the three behaviors rated as most distracting when driving were (i) writing text messages (41%), (ii) reading text messages (62%), and (iii) using a cellular telephone handheld (52%). Victor [3] found that it can lead to that the vehicle frequently deviate lanes when the drivers eyes off the road, even lead to the steering wheel is out of control or the brake is not timely. "Moving object in vehicle" and "Outside person, object or event" could lead to the drivers eyes off the road for a long time which takes a higher risk. The present study has consistency with the foreign research.

Driver Personal Characteristics in the Frequency of Driver Distractions

Driver personal characteristics can affect the frequency of driver distractions. In this paper, the frequency of driver distractions was studied in the aspects of gender, profession, age, and driver experience.

Driver Distraction by Gender and Profession

Independent sample *T*-test was undertaken to investigate the difference between different genders and profession groups of driver in the frequency of 13 kinds of driver distractions. The results are shown in Table 5.

According to the results of *T*-test, combined with average analysis in terms of gender, male and female drivers have no obvious difference in driver distractions in addition to smoking. Results have shown that frequency of male drivers in "Smoking" was a bit higher than the female. In previous observation research [14, 15], there were no significant differences by gender. But Astrain et al. [21] found a

Table 5 Independent sample *T*-test results based on the gender and profession

	1	2	3	4	5	6	7	8	9	10	11	12	13
Sig level (gender)	0.10	0.91	0.29	0.31	0.60	0.18	0.00	0.27	0.85	1.00	0.09	0.70	0.06
Sig level (profession)	0.00	0.09	0.12	0.02	0.15	0.00	0.00	0.36	0.11	0.76	0.07	0.22	0.15

significantly higher rate of mobile phone usage by males. In this present study, in addition to smoking, others were no significant difference by gender. Therefore, it appears that there are no reliable gender differences in driver distractions while driving.

In terms of profession, professional driver with "1. answer the phone" and "7. smoking" occurred more frequent than non-professional driver, but the frequency of "4. Chat" and "6. Operating equipment" is low. In the previous research, the study of driver distractions in dangerous goods drivers is rare. Hanwoski et al. [18] used a naturalistic approach to study the driver distraction of truck driver: 178 distraction-related critical incidents were analyzed, and 34 unique distraction types were identified, but not the types of distractions. This finding shows that professional drivers are more focused on driving, unless they need to answer the call for the transport task.

Driver Distraction by Age

According to age, these respondents were divided into four groups: <25, 25–40, 40–50, and >50. One-way ANOVA method was undertaken to investigate the difference between different age groups of driver in the frequency of 13 kinds of driver distractions. The results are shown in Table 6.

From Table 6, under the condition of the confidence interval of 95%, different age groups of drivers existed significant differences in the frequency of six driver distractions, in which the serial numbers were 4, 6, 8, 9, 12, and 13. Taken "4. Chat," for example, by multiple comparison analysis, the results are shown in Table 7.

The <25 group and 40–50, >50 groups existed significant differences in the frequency of "4. Chat," which can be seen from Table 7. The results is same to "the 25–40 group" VS "the 40–50, >50 group," and "the 40–50 group" VS ">50 group." (VS represents significant differences between different groups). The same as the analysis process of "4. Chat," multiple comparison was used to analyze these driver distractions, in which the serial numbers were 1, 6, 8, 9, 12, and 13 combined with the mean analysis (average in brackets in the following table). The results are shown in Table 8.

The results shown, in the age groups younger and middle-aged drivers (<40), have more frequency than middle-aged and older driver (>40) in no 4, no 6, no 8,

Table 6 Results of one-way ANOVA method in different groups

	1	2	3	4	5	6	7	8	9	10	11	12	13
F	2.699	1.762	0.964	6.845	1.048	3.173	2.565	3.912	4.184	1.143	2.053	4.747	6.509
Sig	0.056	0.154	0.410	0.000	0.372	0.025	0.055	0.009	0.006	0.332	0.107	0.003	0.000

Table 7 Results of multiple comparison analysis in "4. Chat"

Dependent variable			Mean difference $(I - J)$	Std. error	Sig.
4	<25	25–40	0.125	0.180	0.487
		40–50	0.725[a]	0.213	0.001
		>50	0.694[a]	0.259	0.008
	25–40	<25	−0.125	0.180	0.487
		40–50	0.600[a]	0.164	0.000
		>50	0.569[a]	0.220	0.010
	40–50	<25	−0.725[a]	0.213	0.001
		25–40	−0.600[a]	0.164	0.000
		>50	−0.031	0.249	0.902
	>50	<25	−0.694[a]	0.259	0.008
		25–40	−0.569[a]	0.220	0.010
		40–50	0.031	0.249	0.902

[a]Existed significant difference, 0.05

Table 8 Results of multiple comparison analysis in six driver distractions

Driver distraction	Multiple comparison results	
4	<25 (3.59) **VS** 40–50 (2.86), >50 (2.89)	25–40 (3.46) **VS** 40–50 (2.86), >50 (2.89)
6	<25 (3.02) **VS** 40–50 (2.52), >50 (2.46)	25–40 (2.95) **VS** 40–50 (2.52), >50 (2.46)
8	<25 (2.04) **VS** 40–50 (1.64), >50 (1.5)	25–40 (1.96) **VS** 40–50 (1.64), >50 (1.5)
9	<25 (2.72) **VS** 40–50 (2.03)	25–40 (2.62) **VS** 40–50 (2.03)
12	<25 (2.13) **VS** 25–40 (1.72), 40–50 (1.62), >50 (1.57)	
13	<25(2.5) **VS** 25–40 (2.09), 40–50 (1.76), >50 (1.79)	

VS represents significant differences between different groups

no 9, the four driver distractions. Younger drivers (<25) were more likely, than middle-aged and older drivers, to be observed "Moving object in vehicle," "Outside person, object or event." The results indicated that younger drivers prefer to talk with others and easy to be distracted, and the older drivers are stable and do not often use on-board equipment. The results were consistent with other related research. This finding was in partial agreement with these previous researches. Sullman [14] found younger drivers were more likely to be observed engaged in a secondary task while driving. McEvoy et al. [23] found that young drivers (those aged 17–29) reported more crashes which involved a distracting activity than did older drivers, which shows a clear trend for younger drivers to engage more frequency in a number of types of distracted driving. Prat [16] found that younger drivers were more likely, than older drivers, to be observed using a mobile phone,

texting or keying numbers, and drinking. Therefore, in general, the findings of the current study are consistent with previous national and international research, which shows a clear trend for younger drivers to engage more frequently in a number of types of distracted driving. This again highlights the need for targeting younger drivers with interventions aimed at reducing this type of dangerous driving.

Driver Distraction by Driving Experience

According to driver age, these respondents were divided into four groups: <5 years, 5–10 years, 10–20 years, >20 years. One-way ANOVA method was undertaken to investigate the difference between different driver age groups of driver in the frequency of 13 kinds of driver distractions. The results are shown in Table 9.

From Table 9, under the condition of the confidence interval of 95%, different driver age group drivers existed significant differences in the frequency of two driver distractions in which serial number was 1 and 9. Taken "1. Answer the phone," for example, multiple comparison was used to analyze, and the results are shown in Table 10.

From the above table, the <5 years group and 5–10, 10–20 years group existed significant differences in the frequency of "1. Answer the phone." The results is same to the <5 years VS 10–20, >20 years group, the 5–10 years group VS 10–20, >20 years group. (VS represent significant differences between different groups). The same as the analysis process of "1. Answer the phone," multiple comparison was used to analyze "9. Use the navigation," combined with the mean analysis (average in brackets in the following table). The results are shown in Table 11.

From the table, two driver distractions have relationship with driver experience, which were "Answer the phone" and "Use the navigation." Data show that novice drivers have less answer the phone than experienced drivers, which could be due to novice driver was so cautious too much that the driver distraction "Answer the phone" was less in driving process. But the novice drivers in the number of "Use the navigation" than the experience drivers, which could be the novice drivers relying more on vehicle navigation. Fitch et al. [22] found distracted novices at much higher crash and near-crash risk than experienced drivers based on crash and near-crash data, which also suggests that novice drivers need to be more focused on driving.

The Classification of Driver Distractions

To understand the influence of 13 driver distractions on the driving operations, clustering analysis was used to classify the 13 driver distractions based on "direction control ability," "brake response ability," "road sign recognition ability," "traffic forecasting ability," respectively. Taken "direction control ability," for

Table 9 Results of one-way ANOVA method in different groups

	1	2	3	4	5	6	7	8	9	10	11	12	13
F	5.698	0.326	0.444	4.175	1.247	0.868	0.414	1.112	5.270	1.062	0.844	1.113	2.487
Sig	0.001	0.806	0.722	0.006	0.293	0.458	0.743	0.344	0.001	0.365	0.471	0.344	0.061

Table 10 Results of multiple comparison analysis in "1. Answer the phone"

Dependent variable			Mean difference $(I - J)$	Std. error	Sig.
1	<5	5–10	−0.469[a]	0.149	0.002
		10–20	−0.664[a]	0.191	0.001
		>20	−0.186	0.369	0.615
	5–10	<5	0.469[a]	0.149	0.002
		10–20	−0.195	0.206	0.344
		>20	0.283	0.377	0.454
	10–20	<5	0.664[a]	0.191	0.001
		5–10	0.195	0.206	0.344
		>20	0.478	0.396	0.228
	>20	<5	0.186	0.369	0.615
		5–10	−0.283	0.377	0.454
		10–20	−0.478	0.396	0.228

[a]Existed significant difference, 0.05

Table 11 Results of multiple comparison analysis in two driver distractions

Driver distraction	Multiple comparison results
1	<5 years (2.31) VS 5–10 (2.78), 10–20 years (2.98)
9	<5 years (2.71) VS 10–20 (2.02), >20 years (1.8)
	5–10 years (2.48) VS 10–20 (2.02), >20 years (1.8)

example, the clustering result of secondary tasks based on directional control is shown in Fig. 3.

Combined with the frequency distribution of secondary tasks based on directional control (see Fig. 4). The 13 types of driver distractions could be classified to three categories based on the indicator "direction control ability": the low level, the middle level, the high level. Analysis procedure of the other three indicators was similar to the indicator "direction control ability." The result of classification is shown in Table 12.

From the results of classification, most drivers considered the first categories ("3. Listen to music," "4. Chat," and "5. Listen to the radio") had low influence on each driving operations; the second categories ("6. Operating equipment," "7. Smoking," "8. Eating or Drinking," "9. Use the navigation," and "11. Still garbage") had middle influence; and the third categories ("1. Answer the phone," "2. Edit text," "10. Daydreaming," "12. Moving object in vehicle," and "13. Outside person, object or event") had high influence.

According to the multiple resource theory of Wickens [24], the first category has low influence, where these driver distractions do not need to take up visual and action resources, and only take part of the cognitive resources. So these driver distractions and the main driving task do not exist fierce competition, which means drivers do not need too much attention about the first kind of driver distractions.
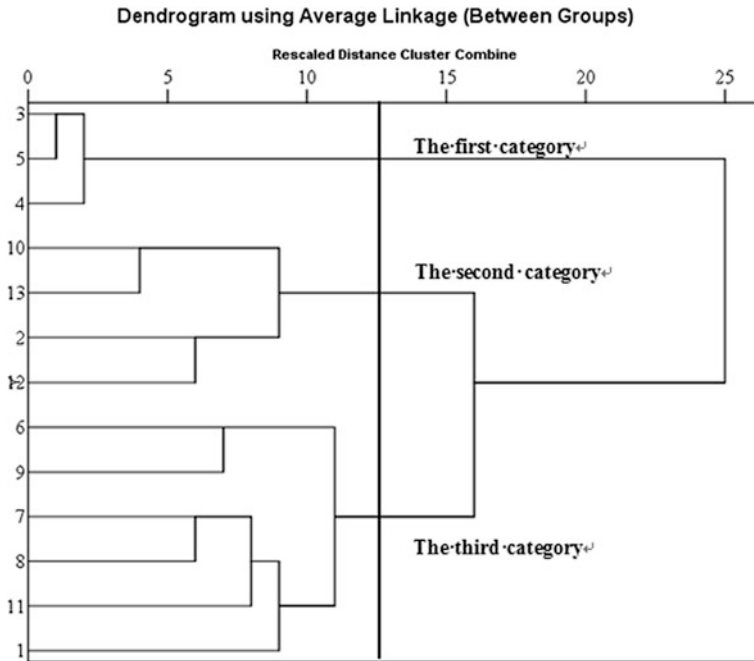
Fig. 3 Clustering results of secondary tasks based on directional control

Fig. 4 Frequency distribution of secondary tasks based on directional control
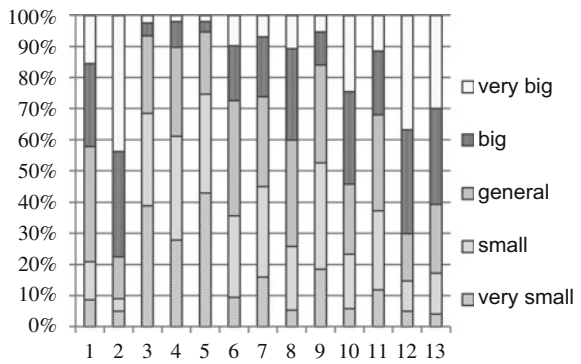
Table 12 Classification of secondary tasks based on four indicators

	Direction control ability	Brake response ability	Road sign recognition ability	Traffic forecasting ability
First	3, 4, 5	3, 4, 5, 9	3, 4, 5	3, 4, 5
Second	1, 6, 7, 8, 9, 11	6, 7, 8, 9, 11	6, 7, 8, 9, 11	6, 7, 8, 9, 11
Third	2, 10, 12, 13	1, 2, 10, 12, 13	1, 2, 10, 12, 13	1, 2, 10, 12, 13

The second category has middle influence which only occupies part of resources and less visual and cognitive resources. Drivers do not have too many operational actions in the normal driving condition, and these driver distractions have low influence in the normal state. However, when in emergency the driver needs more actions, where this kind of distractions occupies part of the operation resources in the emergency state. If the driver performs this kind of distractions, there will be a fierce competition relationship with the main driving task. The second kind of driver distractions has middle influence, and drivers should reduce its occurrence frequency and correct the bad driving habits. The third category which included "Edit text," "Moving object in vehicle," and "Outside person, object or event" will take up several perception channels at the same time or a large number of resources in one perception channel. These driver distractions and the main driving task existed incentive competition which has high influence on driving operations. "Edit text," for example, has the same demand with the main driving on visual, cognitive, and operation channel, which can form a competitive relationship with the premise that all channels' resources were limited. For the third kind of driver distractions, driver needs to be vigilant and take corresponding preventive measures to reduce the impact on driving safety.

Conclusion

This paper adopted self-report survey to investigate the prevalence and severity of 13 types of driver distractions and analyzed the influences of driver personality traits (i.e., gender, professional, age, and driving experience) on the prevalence of driver distractions. Finally, the clustering analysis method was used to classify driver distractions. The conclusions are as follows:

(1) The most frequent distraction is "Listening to music," and the most dangerous distraction is "Writing text messages."

(2) Three personality traits (professional, age, and driving experience) are found to have a significant impact on the prevalence of driver distractions. The analysis results show professional drivers are more focused on driving, unless they need to answer the call for the transport task. There is a clear trend for younger drivers to engage more frequently in a number of types of distracted driving. Novice drivers have less frequency than experienced drivers on "Answer the phone," but more than the experience drivers on "Use the navigation," which maybe the novice drivers were cautious and more relied on vehicle navigation.

(3) This research presented the classification method of driver distractions. According to the influence of driver distractions on driving operation, these driver distractions can be divided into three categories: low, middle, and high. The first category included three driver distractions: "Listen to music," "Chat" "Listen to the radio"; the second category included five driver distractions: "Operating equipment," "Smoking," "Eating or Drinking," "Use the

navigation," and "Still garbage"; the third category included five driver distractions: "Answer the phone," "Edit text," "Daydreaming," "Moving object in vehicle," "Outside person, object or event."

(4) The multiple resource theory was used to explain the rationality of the three categories in the view of cognition and gives some suggestions to drivers. Driver does not need too much attention about the first kind of driver distractions. For the second kind of driver distractions, driver should reduce the occurrence frequency and correct the bad driving habits. For the third kind of driver distractions, driver needs to be vigilant and take corresponding preventive measures to reduce the impact on driving safety.

References

1. Sullman, M.J.M. 2012. An observational study of driver distraction in England. *Transportation Research Part F: Traffic Psychology and Behaviour* 15 (3): 272–278.
2. Horrey, W.J., and D.J. Simons. 2007. Examining cognitive interference and adaptive safety behaviours in tactical vehicle control. *Ergonomics* 50 (8): 1340–1350.
3. Victor, T. 2005. *Keeping Eye and Mind on the Road*. Department of Psychology.
4. Liang, Y., and J.D. Lee. 2010. Combining cognitive and visual distraction: Less than the sum of its parts. *Accident; Analysis and Prevention* 42 (3): 881–890.
5. Papantoniou, P., E. Papadimitriou, and G. Yannis. 2015. Assessment of driving simulator studies on driver distraction. *Advances in Transportation Studies* 35: 129–144.
6. Liang, Y., Horrey, W.J., & Hoffman, J.D. 2015. Reading text while driving: understanding drivers' strategic and tactical adaptation to distraction. *Human Factors the Journal of the Human Factors & Ergonomics Society* 57 (2).
7. McEvoy, S., & Stevenson, M. (2008). Measuring exposure to driver distraction. In *Driver Distraction: Theory, Effects and Mitigation* eds. M.A. Regan, J.D. Lee, & K.L. Young 73–83. Boca Raton, FL: CRC.
8. Royal, D. 2002. National survey of distracted and drowsy driving attitudes and behaviors: 2002: volume 1: findings report. *Indonesian Journal of Animal & Veterinary Sciences*.
9. Mcevoy, S.P., M.R. Stevenson, and M. Woodward. 2006. The impact of driver distraction on road safety: Results from a representative survey in two Australian states. *Injury Prevention* 12 (4): 242–247.
10. RAC Motor Insurance. 2009. The 9 worst driver distractions.
11. Lansdown, T.C. 2012. Individual differences and propensity to engage with in-vehicle distractions—A self-report survey. *Transportation Research Part F Traffic Psychology & Behaviour* 15 (1): 1–8.
12. Taylor, D.M.D., C.E. MacBean, A. Das, and R.M. Rosli. 2007. Handheld mobile telephone use among Melbourne drivers. *Medical Journal of Australia* 187 (8): 432–434.
13. Narine, S., Walter, L.K., & Charman, S.C. 2010. Mobile phone and seat belt usage rates in London 2009. Trl Published Project Report.
14. Sullman, M.J.M. 2010. An observational survey of driving distractions in England. Human factors in road and rail transport: 287–295.
15. Gras, M.E., Planes, M., Font-Mayolas, S. et al. 2012. Driving distractions in Spain. In *Human Factors in Road and Rail Transport ed.* Dorn, L, 299–305.
16. Prat, F., Planes, M., Gras, M.E. et al. 2015. An observational study of driving distractions on urban roads in Spain. *Accident Analysis & Prevention* 74c: 8–16.

17. Stutts, J., J. Feaganes, D. Reinfurt, et al. 2005. Driver's exposure to distractions in their natural driving environment[J]. *Accident Analysis and Prevention* 37 (6): 1093–1101.
18. Hanowski, R.J., M.A. Perez, and T.A. Dingus. 2005. Driver distraction in long-haul truck drivers. *Transportation Research Part F Traffic Psychology & Behaviour* 8 (6): 441–458.
19. Stutts, J.C., D.W. Reinfurt., L. Staplin, and E.A. Rodgman. 2001. *The Role of Driver Distraction in Traffic Crashes. AAA Found.* Traffic Safety, Washington, DC, Tech. Rep..
20. Dot, H.S. 2010. An Analysis of Driver Inattention Using a Case-Crossover Approach on 100 Car Data: Final Report. Drowsiness.
21. Astrain, I., J. Bernaus, J. Claverol, et al. 2003. Prevalence of mobile phone use while driving vehicles. *Gaceta Sanitaria* 17 (1): 59–65.
22. Gregory M, Fitch., Susan A, Soccolich., & Feng, Guo, et al. 2013. The impact of hand-held and hands-free cell phone use on driving performance and safety-critical event risk. Driver Performance.
23. Mcevoy, S.P., M.R. Stevenson, and M. Woodward. 2007. The prevalence of, and factors associated with, serious crashes involving a distracting activity. *Accident Analysis and Prevention* 39 (3): 475–482.
24. Wickens, C.D. 1991. Processing resources and attention. In *Multiple Task Performance* ed. Damos D. London: Taylor and Francis.

Origin-Destination Distribution Prediction Model for Public Bicycles Based on Rental Characteristics

Shuichao Zhang, Yanjie Ji, Dong Sheng and Jibiao Zhou

Abstract Accurate prediction of the rental demand origin-destination distribution of public bicycles provides a foundation according to which layout planning, operational management and dispatching of bicycle sharing system stations may be achieved. Based on the conventional double-constrained gravity model, the rental duration distribution function was employed as a distribution impedance function in order to establish a prediction model for the origin-destination distribution of public bicycles in a bicycle sharing system. The expense incurred by the weighted average travel time of the bicycle sharing system located in the old town of Zhenhai District, Ningbo, was applied to test the origin-destination distribution prediction model for public bicycles based on characteristics of rental duration distribution. Results indicate that the established model demonstrates high precision and can be used to effectively predict the origin-destination distribution of bicycle sharing systems, thus avoiding the dense distribution over short distances which results from the conventional double-constrained gravity model.

Keywords Urban transport · Public bicycles · Rental duration · Origin-destination distribution prediction

Introduction

Compared to private bicycles, public bicycles have been well received by urban residents due to their convenience of rental and travel [1]. Meanwhile, the government has strongly advocated travel by public transport and invested heavily in the construction of the public bicycle system in order to ease urban traffic con-

S. Zhang (✉) · D. Sheng · J. Zhou
School of Civil and Transportation Engineering, Ningbo University of Technology, Ningbo 315211, China
e-mail: zhangsc2588509@126.com

S. Zhang · Y. Ji
Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing 210096, China

gestion. In recent years, the public bicycle system in China has developed rapidly, which has been made available in 123 cities in just a few years, with 510,000 public bicycles and 20,000 rental stations in operation. This has made China a country in the world that owns the most public bicycles [2]. However, as an emerging public transport system, the public bicycle system in most cities has a lot of problems concerning station layout, stand allocation, vehicle dispatch, which greatly affect its overall benefits. In addition, a lack of relevant fundamental theories is also one of the major concerns that need to be addressed during the development of the public bicycle system [3]. The origin-destination (OD) distribution prediction for public bicycles studied in this paper is an important element of the demand prediction for a public bicycle system, as well as an important theoretical foundation for its station layout, vehicle allocation and operation and dispatch and is of both theoretical and practical significance.

Overall, there have been relatively few studies on the OD distribution prediction for a public bicycle system at home and abroad. Domestic studies on a public bicycle system have focused primarily on the status quo of development of public bicycles and their development countermeasures [4, 5], the station layout of the public bicycle system [6, 7], and the connection between rail transit stations and the public bicycle system [8], but hardly on the OD distribution prediction for a public bicycle system proposed in this paper. Wu Yao established a rental demand prediction model for public bicycles by combining the future annual resident travel OD and the modal share of public bicycles based on the multinomial logit model [9]. Liu [6] predicted the number of public bicycles at any rental station when a dispatched bicycle travels from the current location to such station, based on the historical travel OD and the public bicycle rental and return information for previous periods for various rental stations and established a short-term multi-period demand prediction model for rental stations. Based on studies on the methods for location of public bicycle rental stations, Li Tingting and Luo Haixing analyzed and proposed the factors that influence travel mode choice by beginning with travelers' destination choice behavior and established a public bicycle traffic flow prediction model using the multiple logit model [10, 11]. The above analysis indicates that current studies on public bicycle demand prediction have primarily used overall urban OD distribution data and then used modal share to determine the OD data for public bicycles. This method is easy to use but has low-prediction accuracy and cannot reflect the great demand for public bicycles in terms of short-distance trips.

Due to the great differences in the urban traffic environment and the public bicycle system size between foreign and domestic cities, relevant foreign studies have focused primarily on the policies on public bicycle development and its safety [12], but hardly on public bicycle demand prediction. Inês Fradea studied the factors that influence the rental demand for public bicycles in Coimbra, Portugal, and obtained the distribution characteristics of rental demand for public bicycles in Coimbra [13]. After analyzing the public bicycle rental data for 38 cities in relevant countries and areas, Oliver O'Brien pointed out that the public bicycle system is a relatively closed system that can be modeled and analyzed [14].

In view of current studies on public bicycle systems at home and abroad, this paper will improve the doubly constrained gravity model based on the public bicycle rental duration distribution function and establish an OD distribution prediction model that is in line with the demand characteristics of public bicycle systems in order to provide a theoretical basis for the planning, construction and operations management of public bicycle systems.

OD Distribution Prediction Model

The conventional OD distribution prediction models in the field of traffic engineering primarily include gravity models and growth factor models, and the doubly constrained gravity model has been most widely used [15]. In the following, an introduction was given to the OD distribution prediction model for public bicycle systems based on the doubly constrained gravity model.

General Form of the Doubly Constrained Gravity Model

The doubly constrained gravity model uses the law of universal gravitation in physics to describe the problem of traffic attraction between OD zones in a traffic system. It assumes that the number of trips from traffic zone i to traffic zone j is proportional to the number of trip productions at traffic zone i and the number of trip attractions at traffic zone j and is inversely proportional to the impedance to travel between traffic zones i and j. This model often takes the following form:

$$Q_{ij} = K_i K_j O_i D_j f(T_{ij}) \tag{1}$$

$$K_i = \frac{1}{\sum_j K_j D_j f(T_{ij})} \tag{2}$$

$$K_j = \frac{1}{\sum_i K_i O_i f(T_{ij})} \tag{3}$$

In this equation, Q_{ij} represents the number of trips from traffic zone i to traffic zone j, or the number of trips from station i to station j for a public bicycle system; T_{ij} represents the temporal or spatial distance between station i and station j; $f(T_{ij})$ represents the travel impedance function; and K_i and K_j are parameters to be determined.

When $f(T_{ij})$ is known, K_i can be first calculated by letting $K_j = 1$, and then be used to calculate K_j. Equations (2) and (3) are solved iteratively until they converge. The iterative formulae are:

$$K_i^{(m)} = \frac{1}{\sum\limits_j K_j^{(m)} D_j f(T_{ij})} \tag{4}$$

$$K_j^{(m)} = \frac{1}{\sum\limits_i K_i^{(m)} O_i f(T_{ij})} \tag{5}$$

Distribution Impedance Function

Public bicycle rental duration can be expressed by the temporal cycling distance between an origin station and a destination station, which typically reflects the spatial cycling distance between the two stations. The impedance function of travel between two stations actually reflects a cyclist's willingness to travel between the two stations. In reality, the functional expression of rental frequency depending on rental duration can be used to reflect changes in a cyclist's willingness to travel with the temporal distance between two stations, namely the travel impedance function.

Since a bicycle is human-powered, different people cycles at almost the same rate. In addition, longer cycling requires more physical strength and gains less advantage. Existing studies have revealed that the distribution characteristics of public bicycle rental duration for different cities are extremely similar, as shown in Fig. 1 [16].
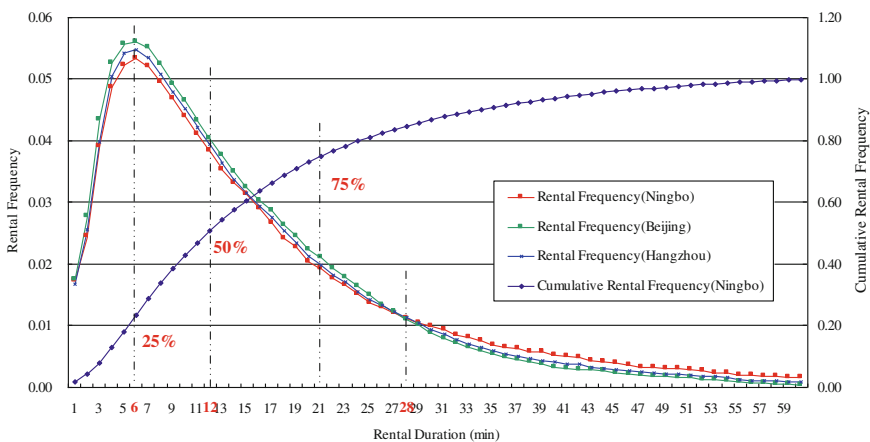
Fig. 1 Analysis of the characteristics of public bicycle rental duration for three cities

The distribution of public bicycle rental duration also satisfies the following functional expression [16]:

$$F(T) = \begin{cases} C_1 T e^{-k_1 T}, & T \le T_0 \\ C_2 e^{-k_2 T} & T > T_0 \end{cases} \tag{6}$$

In this equation, T represents rental duration; $F(T)$ represents the corresponding rental frequency, meaning that $f(T_{ij})$ can be taken as $F(T_{ij})$; and C_1, C_2, k_1, k_2, and T_0 are all parameters which can be determined by fitting using the public bicycle rental data for real cities. For example, an analysis of the public bicycle data for Ningbo, Hangzhou, and Beijing led to the concrete expressions of $F(T)$ for the three cities, as shown in Table 1 [16].

Algorithm Design

Based on the OD distribution prediction model established above, the following algorithm can be used for calculation:

(1) Obtain the characteristic function of rental duration based on the operational data for a public bicycle system, and then obtain the travel impedance function of the public bicycle system, $f(T_{ij})$;
(2) Let $m = 0$, where m represents the number of operations;
(3) Calculate $K_i^{(0)}$ by letting $K_j^{(0)} = 1$;
(4) Calculate $K_j^{(m+1)}$ based on $K_i^{(m)}$, and then calculate $K_i^{(m+1)}$ based on $K_j^{(m+1)}$;
(5) Finish the operation when $\left(K_i^{(m)} - K_i^{(m-1)}\right)/K_i^{(m-1)} \le 0.05$ and $\left(K_j^{(m)} - K_j^{(m-1)}\right)/K_j^{(m-1)} \le 0.05$, otherwise go to step (4) and repeat the process;
(6) Calculate the entire OD matrix, $\{Q_{ij}\}$, after calculating the final $K_i^{(m)}$ and $K_j^{(m)}$.

Model Testing

The OD distribution prediction model established above was tested against the public bicycle system in the old town in Zhenhai District, Ningbo. The old town in Zhenhai District, which is located in the northeast corner of Ningbo and covers an area of about 5.0 km^2, now has 50 pubic bicycle stations, with about 1000 public bicycles available, as shown in Fig. 2. Due to geographical limitations, the distances between the public bicycle stations inside and outside the old town in Zhenhai District are all longer than 5 km. Therefore, there are few connections between the public bicycle stations inside and outside the old town, and the public

Table 1 Characteristic functions of public bicycle rental duration for three cities (T measured in min)

City	Ningbo	Beijing	Hangzhou
Model function	$F(T) = \begin{cases} 0.02437Te^{-0.1667T}, & T \leq 17 \\ 0.0706e^{-0.0641T}, & T > 17 \end{cases}$	$F(T) = \begin{cases} 0.0254Te^{-0.1667T}, & T \leq 17 \\ 0.1558e^{-0.0948T}, & T > 17 \end{cases}$	$F(T) = \begin{cases} 0.0248Te^{-0.1667T}, & T \leq 17 \\ 0.0994e^{-0.0784T}, & T > 17 \end{cases}$

Fig. 2 Distribution of public bicycle stations in the old town in Zhenhai district

bicycle system in the old town is relatively self-contained. The old town has been well developed, and its public bicycle system has operated stably, with about 5.5 rentals per public bicycle. Currently, the old town has 50 public bicycle stations, which size is suitable for computation and testing for public bicycle OD distribution prediction using a computer.

The public bicycle rental data for the old town in Zhenhai District for a number of weekdays and the public bicycle OD distribution data for trips between stations in the old town for typical weekdays were obtained, as shown in Tables 2 and 3.

Table 2 Present OD matrix

	1	2	3	4	5	6	7	8	9	…	49	50
1	12	2	0	9	11	1	0	0	1	…	7	3
2	1	12	1	12	8	0	0	2	2	…	9	3
3	1	3	18	5	6	1	0	0	6	…	10	22
4	6	8	10	10	5	1	0	1	6	…	4	7
5	7	7	10	3	11	4	4	2	7	…	5	15
6	1	0	0	2	3	3	0	1	1	…	2	1
7	0	0	1	1	2	0	6	0	0	…	1	1
8	0	1	0	3	2	1	0	6	0	…	0	3
9	3	3	3	2	4	1	0	2	13	…	15	10
…	…	…	…	…	…	…	…	…	…	…	…	…
49	7	11	7	7	9	1	0	0	6	…	14	6
50	1	4	32	13	15	2	0	2	6	…	6	14

Table 3 Number of trip productions and trip attractions at various stations

Station no.	Number of trip productions	Number of trip attractions	Station no.	Number of trip productions	Number of trip attractions	Station no.	Number of trip productions	Number of trip attractions
1	73	67	18	78	81	35	69	73
2	87	88	19	112	126	36	190	188
3	182	190	20	141	141	37	80	86
4	184	160	21	175	173	38	41	53
5	185	187	22	162	162	39	112	107
6	51	54	23	122	118	40	115	109
7	24	24	24	124	128	41	120	123
8	49	45	25	173	175	42	102	102
9	117	98	26	118	131	43	26	30
10	85	81	27	148	154	44	35	41
11	140	149	28	78	80	45	30	29
12	245	234	29	94	94	46	9	19
13	107	106	30	84	67	47	53	51
14	217	221	31	126	115	48	34	40
15	109	108	32	44	47	49	68	72
16	216	201	33	78	58	50	19	28
17	203	200	34	117	137	Total	5351	5351

The average of the actual rental duration data for three weekdays was used as the temporal distance between the stations, as shown in Table 4.

Based on the data contained in Tables 3 and 4, the OD distribution prediction model established in this paper was used for iterative computation to obtain the predicted OD matrix, as shown in Table 5.

Table 4 Temporal distances between stations

D O	1	2	3	4	5	6	7	8	9	…	49	50
1	18	2	3	4	6	6	7	7	9	…	42	42
2	2	22	2	2	3	4	5	6	8	…	43	45
3	3	2	17	2	3	3	4	5	7	…	42	42
4	4	2	2	24	2	3	4	3	4	…	39	37
5	6	3	3	2	21	3	3	3	4	…	38	39
6	6	4	3	3	3	17	2	3	3	…	38	36
7	7	5	4	4	3	2	15	2	2	…	35	34
8	7	6	5	3	3	3	2	25	2	…	36	33
9	9	8	7	4	4	3	2	2	23	…	34	33
…	…	…	…	…	…	…	…	…	…	…	…	…
49	42	43	42	39	38	38	35	36	34	…	24	6
50	42	45	42	37	39	36	34	33	33	…	6	21

Table 5 Predicted OD matrix

D O	1	2	3	4	5	6	7	8	9	…	49	50
1	11	1	2	7	9	4	2	3	4	…	1	2
2	3	13	3	11	9	2	2	1	3	…	12	2
3	2	3	15	6	8	2	4	2	8	…	9	23
4	5	7	9	12	6	2	1	2	7	…	5	8
5	8	7	9	3	12	3	5	2	8	…	6	16
6	2	1	1	1	2	3	1	2	1	…	2	2
7	1	2	2	4	2	2	5	4	1	…	2	1
8	1	1	1	2	2	1	2	3	1	…	1	2
9	2	3	4	2	3	2	1	1	9	…	114	11
…	…	…	…	…	…	…	…	…	…	…	…	…
49	8	12	6	8	8	2	3	3	5	…	13	7
50	2	4	31	15	16	1	1	1	5	…	5	15

The weighted average travel cost, R, was introduced in order to test the accuracy of the OD matrix obtained from the model:

$$R = \frac{\sum\limits_{i}\sum\limits_{j} T_{ij}Q_{ij}}{\sum\limits_{i}\sum\limits_{j} Q_{ij}} \qquad (7)$$

The model will be considered accurate if the difference between the weighted average travel cost for the OD matrix obtained from the model, R_2 and that for the present OD matrix, R_1, is smaller than 5%. The weighted average travel cost for the present OD matrix, R_1, was calculated to be 14.88 min and that for the predicted OD matrix, R_2, was calculated to be 15.43 min, with an error of 3.7%. Therefore, the OD distribution prediction model established in this paper proved to satisfy the relevant requirements for public bicycle OD distribution prediction. In addition, Table 5 indicated that there was no excessive distribution of trips between short-distance stations in the predicted OD matrix.

Conclusion

OD distribution demand has conventionally been predicted using the doubly constrained gravity model. The distribution impedance function in this model generally takes the form of a power function, and excessive distribution of trips between short-distance stations is the biggest problem of using a power function as the distribution impedance function.

According to the characteristics of the public bicycle system, the rental duration distribution characteristic function was used as the distribution impedance function based on the conventional doubly constrained gravity model. The distribution characteristic of public bicycle rental duration, which describes the relationship between rental frequency and rental duration obtained from statistical analysis of the rental data for a real city, actually reflects the relationship between the demand for public bicycle rental for travel between two stations and their temporal distance. Therefore, it is particularly suitable to use the rental duration distribution characteristic function as the trip distribution impedance function in public bicycle OD distribution prediction. In this paper, the public bicycle system in a real city was used to test the established public bicycle OD distribution prediction model. The results indicated that the model satisfied the accuracy requirement and also avoided the problem of excessive distribution of trips between short distances that the conventional doubly constrained gravity model has.

The OD distribution prediction model established in this paper is an important element of the demand prediction for the public bicycle system and is expected to provide a theoretical foundation for the station layout, operations management and real-time dispatch of public bicycles.

References

1. Zhu, W., Y.Q. Pang, D. WANG, and X.W. YU. 2012. Travel behavior change after the introduction of public bicycle systems: a case study of Minhang District, Shanghai. *Urban Planning Forum* 5: 76–81.
2. Hang, D. 2015. Current situation and countermeasures of urban public bicycle sharing system. *Traffic & Transportation* 10 (1): 30–31.
3. Qian, J., Z.F. Zheng, and Y.F. Feng. 2010. An assessment of the public bicycle facilities in Hangzhou. *Planners* 26 (1): 71–76.
4. Jiang, W.H., M.M. Yi, and Y. Xu. 2015. Suggestions on the development of urban public bicycle. *Road Traffic and Safety* 15 (2): 61–64.
5. Zhang, Y., and M. Zhao. 2014. Discussion on the efficiency and policy orientation of public bicycle sharing system in China. *Urban Research* 6: 117–123.
6. He, L., D.W. Chen, X.H. Li, and J. Lu. 2012. An optimization model of the layout of public bike rental stations. *Journal of Wuhan University of Technology (Transportation Science & Engineering)* 36 (1): 129–133.
7. Zhou, Y.J. 2012. Urban bicycle sharing system planning. *Urban Transport of China* 10 (5): 50–54.
8. Zhou, Q., G. Wu, and H. Sun. 2015. Characteristics of public bicycle as means of access/egress for metro. *Journal of Transportation Systems Engineering and Information Technology* 15 (3): 179–184.
9. Wu, Y., H. Chen, N. Bao, and W. Feng. 2015. A model development of public bicycle rental demand based on multinomial logit. *Journal of Dalian Jiaotong University* 15 (3): 179–184.
10. Li, T.T. 2010. *Study on locating and planning of urban public bike rental station*. Beijing: Beijing Jiaotong University.
11. Luo, H.X. 2013. Study on Locating and Planning of Urban Public Bike Rental Station. Beijing: Beijing Jiaotong University.
12. Lansell, K. 2011. *Melbourne bike share and public transport integration*. Melbourne: University of Melbourne.
13. Fradea, Inês, and Anabela Ribeiroa. 2014. Bicycle sharing systems demand. *Procedia—Social and Behavioral Sciences* 111: 518–527.
14. O'Brien, Oliver, James Cheshire, and Michael Batty. 2014. Mining bicycle sharing data for generating insights into sustainable transport systems. *Journal of Transport Geography* 34: 262–273.
15. Wang, W., and X.C. Guo. 1998. *Traffic engineer*. Nanjing: Southeast University Press.
16. Zhang, S.C. 2015. Research report of rental characteristics for Ningbo public bicycles system. Ningbo: Ningbo University of Technology.

Study on Selection Methods of Speed Control Measures for Low Grade Roads in Rural–Urban Fringe

Ning Zhong, Yongfeng Ma, Wenjun Zhang and Jian Lu

Abstract With the advancement of mechanization process, road in rural–urban fringe has become an area where traffic accidents happen frequently, and overspeed is the primary cause for fatal traffic accidents. In order to improve the traffic safety level of low-grade roads in rural–urban fringe, speed control measures should be selected scientifically and reasonably. Based on traffic characteristics of low-grade roads in rural–urban fringe and implementation effects of the existing measures, this paper analyzes the applicability of speed control measures under common conditions of low grade roads in rural–urban fringe. Moreover, suggestions are proposed for selection of speed control measures in typical special segments of low grade roads in rural–urban fringe. In addition, by taking the low-grade roads in rural–urban fringe of Lishui City in Zhejiang Province as an example, selection schemes of speed control measures are formulated.

Keywords Rural–urban fringe · Low-grade road · Speed control measure

Introduction

Rural–urban fringe is stipulated as the mixed district planned as construction land, state-owned land and collectively owned land as well as the region planned as agricultural land or included in state-owned construction land [1].

As the intermediate zone of rural–urban transition formed in the interaction between cities and countries, road in rural–urban fringe presents characteristics

N. Zhong · Y. Ma · W. Zhang · J. Lu (✉)
Jiangsu Key Laboratory of Urban ITS, Southeast University,
Nanjing 210096, People's Republic of China
e-mail: lujian_1972@seu.edu.cn

N. Zhong · Y. Ma · W. Zhang · J. Lu
Jiangsu Province Collaborative Innovation Center,
Modern Urban Traffic Technologies, Sipailou #2, Nanjing 210096,
People's Republic of China

different from either urban traffic characteristics or rural traffic characteristics in the aspects of traffic composition and speed distribution. It has the following problems: mutual interference exists among various traffic flows; and imperfections of transportation infrastructures covering motor and non-motor separation [2].

Based on field investigation of county road X603, some critical problems exist in traffic conditions and control situations can be concluded. The educational degree of people in rural–urban area is low and their traffic safety awareness is weak. The most outstanding features of traffic flow and vehicle composition on low grade roads in rural–urban area include severe mixture of transportation, obvious difference of vehicle performance and high degree of speed dispersion. Planning of low grade roads in rural–urban fringe lacks a uniform standard and the linear condition is poor. Topographic relief along the line is huge and the forms of section are different. The speed control effect of low grade roads in rural–urban fringe is not ideal.

With the advancement of mechanization process, road in rural–urban fringe has become an area where traffic accidents happen frequently. The occurrence rate of traffic accidents on road in rural–urban fringe of many cities reaches above 50% [3, 4]. This paper mainly studies low grade roads in rural–urban fringe, i.e., road with the executive rank of county level or below and the technical grade of not exceeding level 2, or classified road with the technical grade of not exceeding level 3. Though the flow is lower than that of high-grade roads, the occurrence rate of traffic accidents stays at a high level. Among fatal traffic accidents that happen on low grade roads in rural–urban fringe, excessive speed is an important reason that should not be ignored.

At present, most domestic and overseas studies on speed control focus on high-grade highways, including strengthening of road monitoring, reinforcement of pavement management, and setting of traffic signs for traffic management. Besides, speed control measures were evaluated at home and abroad. Webster indicated that the set of speed hump leads to 71% reduction of traffic accidents. Hirst studied the influence of speed control measurements on traffic accident under different conditions by establishing prediction models in 2005. Based on traffic safety conditions on high-class mountainous highway, road hump was proposed as mandatory speed control measurement on highway, and its effects, security, and reliability were analyzed through simulation [5].

Considering low-grade roads in rural–urban fringe has traffic characteristics different from that of high-grade roads in China, the existing conclusions and achievements aimed at high-grade highways cannot be directly applied to low-grade roads in rural–urban fringe. At present, there is a lack of studies on speed control of low-grade roads in rural–urban fringe.

The objective of this paper mainly focuses on improving traffic safety conditions of low grade roads in rural–urban fringe. Based on the analysis results of applicability according to characteristics of low grade roads in rural–urban area, suggestions about speed control measures on particular sections are proposed in order to achieve traffic safety improvement.

Selection of Speed Control Measures

Applicability of Speed Control Measures

This paper concludes various speed control measures for low-grade roads in rural–urban fringe under conditions of different functional orientations, drivers' characteristics, traffic environments, and construction costs in order to select speed control measures in engineering practice according to the practical situations.

(1) Functional orientation of road

Road functions mainly include maneuverability and accessibility. In view of the basic definition about low-grade roads in rural–urban fringe in this paper, basic types of low-grade roads in rural–urban fringe include: secondary artery, collective highway, and minor road. For secondary artery and primary collective highway dominated by maneuverability, measures with a great influence on driving comfort should not be adopted, such as stone or gravel pavement, humping rubber deceleration strip, concrete prefabricated or cast-in situ deceleration strip and speed hump. For secondary collective highway and minor road dominated by accessibility, speed limitation measures should be selected oriented by speed limitation effect. Measures with good deceleration effect are suggested, such as speed hump and stone or gravel pavement.

(2) Drivers' characteristics

Traffic participants of low-grade roads in rural–urban fringe have weak traffic safety awareness, and the compliance rate is low. Therefore, speed control schemes dominated by mandatory speed control measures and supported by cautionary measures should be adopted in segments where speed must be strictly controlled, such as segments where traffic accidents happen frequently and the access entry of road and principal line, so as to guarantee that drivers will reduce the speed.

Mandatory measures mainly include: humping rubber deceleration strip, concrete prefabricated or cast-in situ deceleration strip, spike deceleration strip, lateral oscillation speed reduction markings, stone (or gravel) pavement, speed hump, speed table, and lane narrowing; cautionary measures mainly include: speed limit sign, thin layer pavement (colorful), speed-monitoring measures, and visual speed reduction marking.

(3) Inducement for traffic safety problems

Overspeed is an important factor of causing traffic safety problems, but it is not the only factor. When speed control measures are selected, the causes and severity degree of traffic safety problems should be analyzed in detail. Moreover, speed control measures with appropriate effect must be selected, so as to make drivers maintain some "sensitivity" toward measures with strong effect. When overspeed is the main cause for black spots, measures with strong effect should be taken;

otherwise, measures without strong effect should be adopted, so as to play a role of warning and caution.

Measures with strong effect mainly include: speed monitoring measures, deceleration strip, speed hump, speed table, and stone (or gravel) pavement.

(4) Length of action zone

Speed control measures of different types have different lengths of action zone. According to the effective scope of measures, measures can be divided into segment type and section type. On low-grade roads in rural–urban fringe, if the speed limitation segment is a continuous accident black spot with poor linear conditions in a long distance, measures of segment type can be considered; measures of section type can be set up for local dangerous points or abrupt change of road cross-section.

Measure of segment type means to control speed within a long segment and to remind drivers of driving cautiously, including interval speed limitation, continuous lateral oscillation speed reduction markings, speed reduction pavement, and longitudinal visual speed reduction markings. In the existing overseas engineering practice, the interval length of segment speed limitation measures is from 2 to 40 km [6]; when the speed limitation segment is shorter than 2 km, this measure is not suggested.

Speed control measure of cross-section type is able to make vehicles reduce speed in the current cross-section and short interval following it. Such measures include speed limit sign, lateral oscillation speed reduction markings, deceleration strip, radar or laser velocimeter, speed feedback mark, traffic circle and traffic bottleneck. When the standard mean square error of spot speed and v_{85} of this point is greater than 10 km/h, speed control measure of cross-section type is suggested [7].

(5) Road cross-section forms

Cross-section forms of road will influence selection and implementation effect of measures to some extent. Low-grade roads in rural–urban fringe mainly includes two-lane and single-lane road.

For single-lane road, owing to the restriction in working principle of equipment, speed monitoring measures and visual speed reduction markings are not suggested. In view of the thoroughfare conditions of single-lane road, lane cross-section narrowing measure is not recommended.

For two-lane road, the influence of medial separator and motor and non-motor separation strip should be considered. When the road has no medial strip, the width of oscillation deceleration strip facility should be the same with the total width of two-lane road. Otherwise, speed limitation measures can be set up in the driving direction of speed limitation only. When mixed traffic flow exists on the road and there is no motor and non-motor separation, space should be reserved for non-motor vehicles and pedestrians.

(6) Construction cost

When speed control measures are selected for low grade roads in rural–urban fringe, economical and effective speed limitation measures should be taken based on the regional economic development level and considering installation, construction and later-stage maintenance cost. Low-cost measures are suggested for low grade roads in rural–urban fringe, and speed limitation measures with relatively high cost can be selected in regions with severe over speed problem and serious hidden danger.

Low-mid cost measures include: speed limit sign, humping rubber deceleration strip, concrete prefabricated or cast-in situ deceleration strip, spike deceleration strip, lateral oscillation speed reduction markings, stone (or gravel) pavement, low-cost thin layer pavement, speed feedback mark, speed hump, speed table, and longitudinal visual speed reduction markings.

High-cost measures include: radar or laser velocimeter, interval speed limitation facilities, traffic circle, pavement narrowing, and illusionary markings.

(7) Traffic characteristics

Vehicle mixture and road access are distinguishing features of low grade roads in rural–urban fringe. When the traffic composition proportion of large vehicles or traffic accident rate caused by large vehicles is high, besides conventional measures, multi-strip and multi-group lateral oscillation speed reduction markings with bright colors, speed hump, and speed table should be adopted for the characteristics of large vehicles such as high viewpoint and big mass. Measures like traffic circle, ripple tracks, and lane narrowing are not suggested. When the traffic volume along the line (including vehicles parked along the road) exceeds more than 10% of traffic volume on the principal line, interval speed limitation is not suggested. When the traffic volume is too high, statistical data effectiveness of this measure will drop obviously.

Suggested Countermeasures of Speed Control on Special Segments

When speed control measures are selected, security, effectiveness, harmony, and economic efficiency of the measure should be considered at the same time. Attention must be paid to the local situations; mechanical application and hyper-correction should be avoided. Traffic characteristics of the road should be analyzed systematically, and speed control measures must be selected scientifically. The installation sites of measures should be determined according to characteristics and requirements of the facilities. Based on typical special segments of low grade roads in rural–urban fringe, this paper proposes some suggestions for speed control measures except speed limit sign.

Segments with Bad Linearity and Segments Crossing Villages and Towns

The low-grade roads in rural–urban fringe have poor technical grade and linear condition, such as abrupt slope, long minus grade, sharp turn, and consecutive curves. Besides installation of corresponding warning signs, speed limitation measures should be taken to make the vehicles maintain a relatively low speed in segments and to guarantee traffic safety of vehicles in segments with bad linearity. Meanwhile, mixture of non-motor vehicles and pedestrians is severe on segments crossing villages and towns in low-grade roads in rural–urban area, and there are many traffic conflicts which might trigger traffic accidents easily. Therefore, running speed of vehicles in segments of villages and towns must be controlled.

Based on different speed limit values and corresponding functional orientations, suggested speed control measures are given for segments with bad linearity and segments crossing villages and towns (Table 1).

Abrupt Change of Road Cross-Section

Due to the differences between road and roadside facilities like bridge and tunnel in planning and construction periods, when the broadened road is connected to narrow bridge or tunnel with the fixed width, abrupt change of road cross-section can be caused easily. In order to avoid traffic accidents, vehicles should maintain a low speed during merging and separation near the abruptly changed cross-section.

Based on different speed limit values and corresponding functional orientations, suggested speed control measures are given under the conditions of abrupt change of road cross-section, as shown in Table 2.

Non-signal Control Intersection and Pedestrian Crosswalk

Non-signal control intersection and pedestrian crosswalk are the concentration zone of traffic conflicts between pedestrians and vehicles, motor and non-motor vehicles, as well as motor vehicles and motor vehicles on low-grade roads in rural–urban fringe. When vehicles pass traffic nodes like intersection without signal control or pedestrian crosswalk, they need to reduce the speed or stop, and continue to go through after confirmation for safety. Therefore, besides warning signs should be set up at the upstream of intersection and pedestrian crosswalk, mandatory speed control measures should be set up based on different speed limit values and corresponding functional orientations (Table 3).

Table 1 Speed control measures under the conditions of segments with bad linearity and segments crossing villages and towns

Speed limit value (km/h)	Interval speed limitation	Visual speed reduction marking	Radar or laser velocimeter	Speed feedback mark	Low-cost thin layer pavement	Stone or gravel pavement	Continuous spike deceleration strip	Continuous lateral oscillation speed reduction marking	Lane section narrowing
20						□	√	√	△
30				△		□	√	√	△
40			△	√			√	√	
50		△	□	√	□		□	√	
60	△	△	√	√	□		□	√	

Note "√" means that the measure is suggested; "□" means that the measure is recommended generally; "△" means that the measure can be applied

Table 2 Speed control measures under the conditions of abrupt change of road cross-section

Speed limit value (km/h)	Spike deceleration strip	Humping rubber deceleration strip	Lateral oscillation speed reduction marking	Speed hump or speed table	Speed feedback mark
20	√	△	□	√	
30	√	△	□	√	
40	√		√	□	△
50	□		√	△	□
60	□		√	△	□

Note "√" means that the measure is suggested; "□" means that the measure is recommended generally; "△" means that the measure can be applied

Table 3 Mandatory speed control measures for intersection without signal control and pedestrian crosswalk

Speed limit value (km/h)	Spike deceleration strip	Humping rubber deceleration strip	Lateral oscillation speed reduction marking	Speed hump or speed table	Traffic circle (applicable to intersection only □)	Lane narrowing
20	√	√	□	√	△	△
30	√	□	□	√	△	△
40	√	△	√	□		
50	□		√	△		
60	□		√	△		

Note "√" means that the measure is suggested; "□" means that the measure is recommended generally; "△" means that the measure can be applied

Regions Near Access Entry

The existence of road access is one of the characteristics of low-grade roads in rural–urban fringe. In regions near access entry, the influence on vehicle traffic of main road should be reduced as far as possible. Therefore, cautionary measures are suggested to remind vehicles on main road of emergencies. On the access road, mandatory measures should be taken to reduce speed, and the principle that main road first must be considered. Access and exit should be completed on the premise of guaranteeing safety (Table 4).

Table 4 Speed control measures for regions near access entry

	Spike deceleration strip	Humping rubber deceleration strip	Lateral oscillation speed reduction marking	Speed hump or speed table	Stone or gravel pavement
Main road	□		√		
Access road	□	√	△	√	△

Note "√" means that the measure is suggested; "□" means that the measure is recommended generally; "△" means that the measure can be applied

Case Study

The county road X603 is located in Suichang County of Lishui City in Zhejiang Province. As a secondary road, the whole line is dominated by mountain landscape. This paper selects Paiqian Village Segment where accidents happen frequently as the speed control implementation case of low-grade roads in rural–urban fringe, to formulate speed control schemes.

Overview About the Segment

Paiqian Village Segment is from 9k + 300 to 10k + 100. Residential and mixed lands are mainly located along the road line, and cultivated and woody lands are along the upstream and downstream lines of the segment. Lane width of this segment is 9 m and the roadbed width is 12.3 m. It is a two-lane road, and a double solid yellow line is used to separate the two lanes. There is no physical isolation between motor and non-motor vehicles or pedestrian lane. Cement concrete pavement is adopted. Design speed of the segment is 60 km/h, and longitudinal slope of the road is smaller than 2%. Traffic accident caused by overspeed happened on this segment during recent 3 years, and the accident happened in the concentration zone of residential area.

Field Survey

Traffic flow of this segment presents a running status of free flow. Table 5 shows the traffic volume and traffic composition during peak hours. Traffic facilities along

Table 5 Traffic volume of the segment during peak hours (unit: set)

Direction	Small car	Medium bus	Large bus	Small truck	Medium truck	Large truck	Agricultural vehicle	Non-motor vehicle
West–east	196	20	0	18	28	6	28	72
East–west	122	26	0	8	22	6	26	42
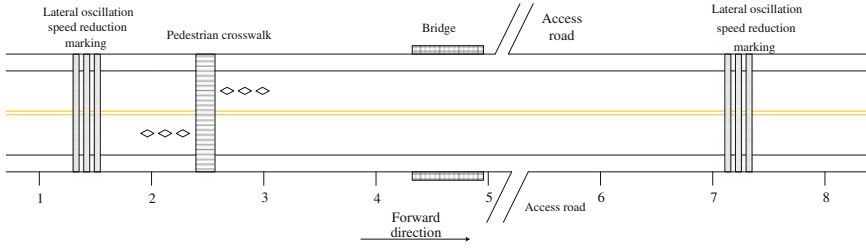
Fig. 1 Schematic diagram for Paiqian Village Segment of X603
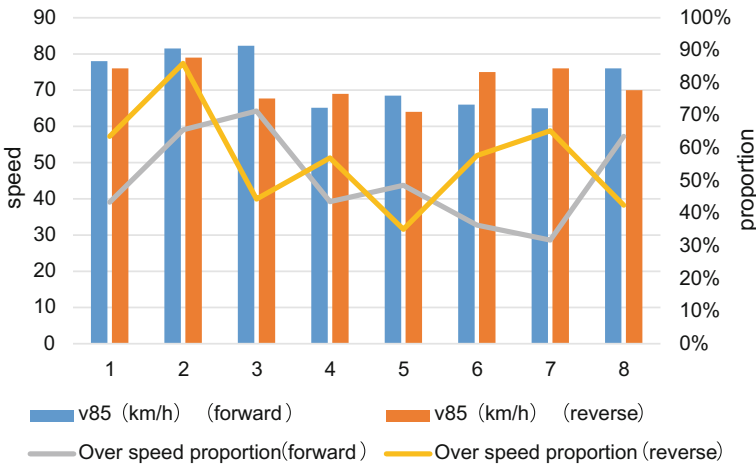
Fig. 2 Schematic diagram for statistics about spot speeds at various observation points (forward driving direction)

the segment line are drawn through spot surveying, and 8 speed observation points are selected (as shown in Fig. 1). The interval between the observation points is 100 m, and 1569 spot speeds are acquired in the survey.

Figure 2 shows the statistical value of spot speeds at various observation points in Paiqian Village Segment as well as statistical value of segment dividing.

Selection of Speed Control Measures

Continuous lateral speed reduction markings have been adopted to control speed in this segment. According to the statistical values of spot speeds, overspeed proportion of the whole segment is high, and the implementation effect of simply relying on lateral oscillation speed reduction markings is not ideal. This road is secondary collective road, and low-mid cost measures with good implementation effect and without influence on road capacity are suggested. Therefore, besides
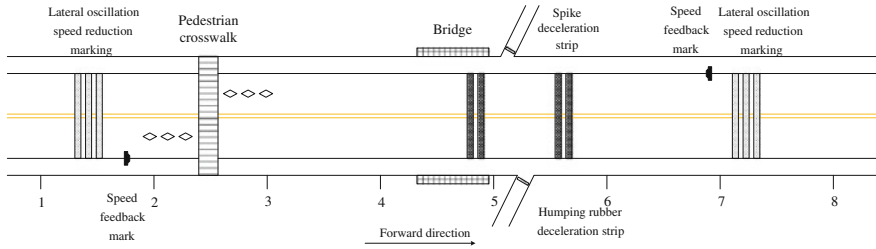
Fig. 3 Schematic diagram for speed control measures for Paiqian Village Segment of X603

lateral oscillation speed reduction markings, speed feedback mark can be set up before entering the village, so as to enhance the speed control effect. Spike deceleration strip should be arranged on the principal line in areas near access entry, and humping rubber deceleration strip or speed table can be set up on the access road. Since there is no median treatment, in order to avoid safety concerns resulted by vehicles changing lanes to keep off deceleration facilities, set the deceleration marks, and speed bumps on two-lane road and reserve space on both sides for pedestrians and non-motor vehicles. The Positions of the facilities are determined according to relevant codes, as shown in Fig. 3.

Conclusion

This paper studied selection methods of speed control measures for low-grade roads in rural–urban fringe. First, traffic characteristics of road and implementation situations of speed control measures were analyzed, and defects of speed control measures for low-grade roads in rural–urban fringe in China were proposed. Besides, application situations of speed control measures under different conditions of low grade roads in rural–urban fringe including functional orientations, drivers' characteristics, traffic environments and construction costs were summarized. In addition, suggestions for speed control measures in special segments of low-grade roads in rural–urban fringe were raised. Finally, based on the practical case, selection schemes of speed control measures for low-grade roads in rural–urban fringe were formulated, so as to provide a reference for traffic safety management of low grade roads in rural–urban fringe.

References

1. Fu, Chengwei, and Mingxing Chen. 2010. Review on urban-rural linkage in China. *Progress in Geography* 29 (12): 1525–1531.
2. Xu, Wangying, Li Xu, and Min Yang. 2008. Study on transportation planning in rural-urban fringe. *Channel Science* (9): 30–31.
3. Guang, Xiaoping, Yongsheng Qian, and Bo Zhou. 2006. Study on road traffic safety design in urban-rural binding region. *Communications Standardization* 5: 131–135.
4. Conggao, Lv, and Wei Yingna. 2007. Analysis and countermeasures of road safety in urban-rural binding regions in China. *Communications Standardization* 12: 180–183.
5. Zhang, Weihua. 2011. *Study of Speed Control Inducement Technology in Mountainous Highway and Wide Area Application*. Xi'an: Chang'an University.
6. Soole, David W., Barry C. Watson, and Judy J. Fleiter. 2013. Effects of average speed enforcement on speed compliance and crashes: A review of the literature. *Accident Analysis and Prevention*, 46–56.
7. Sun, Yingjie. 2008. *Study on Reasonable Speed Limitation Methods for Mountainous Highway Based on Running Speed*. Xi'an: Chang'an University.
8. Li, Yujuan. 2011. *Study of Alignment Optimization Technique for Urban-Rural Fringe Road Based on Traffic Safety*. Wuhan: Wuhan Institute of Technology.
9. Cui, Mengjing. 2014. *Study on Epidemiology of Road Traffic Injury*. Nanjing: Southeast University.
10. Xu, Ting. 2011. *Study on Highway Speed Zone Division and Laddered Transition Section of Speed Limitation*. Beijing: Beijing University of Technology.
11. Institute of Engineers (ITE). 1993. *Document Entitled Speed Zone Guidelines—A Proposed Recommended Practice*. USA: Washington.
12. Wang, Xiaonan. 2009. *Study on Segmentation Methods of Reasonable Speed Limitation for Mountainous Highway*. Xi'an: Chang'an University.
13. Zhong, Xiaoming, Shiwen Zhang, Jia Jia, Jian Zhang, and Hongyuan Wang. 2010. Study on speed management technology. *China Journal of Highway and Transport* 23: 99–104.
14. *Procedures for Establishing Speed Zones*. Texas Department of Transportation, 2006, 11.
15. *Establishment of Speed Limits and Zones*. State of Alask Department of Transportation and Public Facilities, 2000.
16. Manual on Uniform Traffic Control Devices of Australia.
17. Ribbens, H. 1986. *Guidelines for Setting Speed Limits*. South Africa: National Institute for Transport and Road Research, CSIR, 2.
18. Zhao, Yifei. 2009. *Study on Several Problems about Highway Traffic Safety*. Xi'an: Chang'an University.
19. Pan, Binghong. 2008. *Study on Average Longitudinal Slope of Mountainous Highway*. Xi'an: Chang'an University.
20. Sun, Huiyuan, Sun Li, and Wei Ganquan. 2002. Study on relation between highway speed limit and safe sight distance. *Highway* (2): 1–3.
21. Tang, Minwen. 2008. *Study on Highway Speed Limitation Standards and Methods*. Changsha: Hunan University.

Effects of Driver Fatigue and Road Curvature on Steering Wheel Angle

Qingning Niu, Zhiqiang Zhou, Pengcheng Yu, Shuo Liu and Qiuhong Wang

Abstract This paper presents the findings of a simulator study that examined the effects of driver fatigue and road curvature on steering wheel angle. Twenty participants took part in this research, and each participant drove in two sessions, alert driving session and fatigue driving session. Throughout the driving task, the steering wheel angle, roadway scene video, and participants' face video were recorded. According to Karolinska Sleepiness Scale (KSS) score and road curvature, steering wheel angle data were classified into six categories. Three metrics including mean steering wheel angle, standard deviation of steering angle, and coefficient of variation of steering angle were proposed and evaluated to each category. The results revealed that road curvature was the main factor of the change of steering wheel angle and driver fatigue had significant effects on steering wheel angle only on the straight road.

Keywords Driver fatigue · Steering wheel angle · Road curvature

Introduction

Driver fatigue (or driver drowsiness or sleepiness) has been proven to be one of the main causes of traffic accidents [1]. Klauer et al. [2] have shown that fatigue driving increases the accident risk by four to six times, compared with alert driving. In fact, it is found that approximately 20% of all fatal road accidents involve driver fatigue, varying from 5 to 25% in different countries [3]. However, Nordbakke and Sagberg [4] revealed that drivers tend to continue driving, even though they are aware of the risk factors that related to fatigue. In light of these disconcerting statistics,

Q. Niu · Z. Zhou · P. Yu · S. Liu (✉) · Q. Wang
Road Traffic Safety Research Center of Ministry of Public Security, Beijing 100062, China
e-mail: niuqingning@126.com

countermeasures against fatigue driving have received increased attention during the last couple of decades [5].

A change in the mental state can induce a change in driving performance [6, 7]. Otmani et al. [8] found that driver fatigue has a significant effect on driving performance, especially on lane keeping and steering performance [1, 9]. Zhong et al. [10] found that when drivers had a fatigued status, the steering wheel angle and vehicle tracking became irregular, and the range of deviation greatly increased.

Sayed performed a series of experiments on driver's steering characteristics. It has shown that steering activity is a good skill for evaluating driver fatigue among other variables [11, 12]. It has been reported that fatigue drivers have a lower frequency of steering reversals [13], a deterioration of steering performance [14], a decrease in the steering wheel reversing rate [15], more frequent steering maneuvers during wakeful periods and no steering correction for a prolonged period of time followed by a jerky motion during drowsy periods [16], low velocity steering [17], large amplitude steering wheel movements, and large standard deviations in the steering wheel angle [18].

The current research focuses on the correlation between steering-related variables and driver fatigue, while the effects of road curvature are not considered. It is known that road geometry influences steering wheel angle seriously. Weller and Schlag [19] found that the road geometry influences the accuracy of distraction detection based on driving data. Mortazavi et al. [9] found that it was hard to obtain any possible direct correlation between steering-related variables and driver fatigue by analyzing the data from all sections simultaneously. It is necessary to investigate the independent and interaction effects of driver fatigue and road curvature on steering wheel angle. The presented experiment was designed to further investigate the correlation between steering related variables, driving fatigue level and road geometry. Experiments were conducted in a driving simulator. According to the road geometry, road type was classified into three categories. The effects of driver fatigue and road types on steering wheel angle were analyzed, respectively.

Method

Subjects

Twenty participants (6 women and 14 men) aged 25–39 years (mean = 29.85, std = 4.61) took part in this study. All participants were experienced drivers with valid licenses and 3- to 18-year driving experience (mean = 8.05, std = 4.03), who drove at least 12,000 km annually. They were free of any sleep disorder and on no medication, and recruited via an advertisement in school website, each participant was paid for ¥200.

Driving Simulator

The driving task took place in a driving simulator. Road image was placed 2.50 m away from the driver. The driving scenario was a 133-km-long highway of a sampled actual Changping highway located between Changchun and Siping cities, with two lanes in one direction. The selected traffic situation was only sparse oncoming traffic and no traffic driving in the same direction as the test participant.

Experimental Procedure

Each participant took part in two driving simulation sessions, normal driving session (alert driving session) and sleep-deprived session (fatigue driving session). For each participant, it took two days to accomplish the experiment. The two test days were conducted approximately 2 weeks apart, and each participant was informed two days before performing the experiment. Participants were asked to abstain from alcohol and to restrict tea and caffeine consumption 12 h before the test.

For the alert driving session, before the test day, participants were allowed to carry on with their daily life activity and instructed to have at least eight hours of sleep (11:00 p.m.–7:00 a.m.) in the night. In the test day, participants arrived at the laboratory at 8:30 a.m., and experiment started at 9:00 a.m. Before the experiment, participant's personal information was checked and the requirements of the study and details of the protocol were informed. The alert driving session lasted 2 h. In the first 30-min driving, participants completed a practice training session to be familiar with the simulator. The familiar driving data were excluded, and data in the last 90-min driving process were recorded as the alert driving data.

For the fatigue driving session, before the test day, participants were allowed to carry on with their daily life activity and instructed to have only five hours of sleep (1:00 a.m.–6:00 a.m.) in the night. In the test day, participants arrived at the laboratory at 1:30 p.m. The fatigue driving experiments were conducted between 2:00 p.m. and 5:00 p.m. or till the driver was too fatigued to continue driving. Most of the drivers stopped before 5:00 p.m. In this session, participants were sleep-deprived and susceptible to falling asleep during driving.

In the two driving sessions, participants were asked to drive at their own pace to observe the usual driving rules without exceeding the highway motorway speed limit (in China 80–120 km/h). Each participant was questioned about his/her level of fatigue according to the Karolinska Sleepiness Scale (KSS) [20] before and after the driving task. These subjective KSS levels were recorded as the baseline for the researcher to assess participants' fatigue level during the experiment according to the video data.

Recorded Variables

Three types of data were recorded during the experiment: (1) steering performance data including steering wheel angle, which were recorded at 20 Hz. (2) digital video of the roadway scene recorded outside the cabin at 10 Hz. (3) digital video of the driver's face recorded inside the cabin at 10 Hz. All of these data were recorded synchronously.

Data Classification

Face Video Data Classification

Since the way of a driver's facial feature changes as they get progressively fatigue, their facial feature at different fatigue level during the drive can give important insight into his current state [21]. The face video data of each participant were reviewed, and their fatigue level was scored on the KSS by three researchers together. Furthermore, participants' subjective KSS level, which was evaluated by themselves when started and finished the driving experiment, was considered.

Ingre et al. [22] found that the standard deviation of lane position (SDLP) increased gradually from KSS level 1 (SDLP = 0.191, SE = 0.063) to KSS level 3 (SDLP = 0.221, SE = 0.042), while it increased dramatically from KSS level 7 (SDLP = 0.277, SE = 0.042) to KSS level 9 (SDLP = 0.465, SE = 0.040). It indicates that, when KSS level is 3, participant is in alert state and can control vehicle stably, and when KSS level is increased to 7, participant is in fatigue state and the control ability is affected. This conclusion agrees well with previous research on subjective and objective sleepiness [20] and suggests that serious behavioral and physiological changes do not occur until relatively high levels of sleepiness ($KSS \geq 7$) are reached. Meanwhile, based on the face video data, it is impossible for researchers to distinguish participant's fatigue level between a rating of 5 ("neither alert nor sleepy") and a rating of 6 ('some signs of sleepiness'), or between a rating of 5 ("neither alert nor sleepy") and a rating of 4 ('rather alert'). However, there are significant differences between a rating of 3 ('alert') and a rating of 7 ('sleepy, no effort to stay awake'). In this research, if $KSS \leq 3$, driver is in alert state, if $KSS \geq 7$, driver is in fatigue state, and other KSS levels are not discussed.

The driving session was coded and cataloged according to the KSS level, which could make it easy to locate notable points in the data. All datasets were reviewed and coded by the same researchers. The coded episodes information includes: subject number, KSS level, start time of the driving state, end time of the driving state, and comments.

Road Video Data Classification

The steering wheel angle values along the road were dependent upon road curvature [9]. To compare steering wheel angle data for different driving states under different road curvatures, the driving performance data should be examined under different road geometry, i.e., straight road, curvature road, and connection road (connection between straight road and curvature road). The road video data of each participant were reviewed based on the start time and the end time of the divided face video episode. According to road geometry, road video data were divided into three types of episodes: straight road episode, curvature road episode, and connection road episode.

Steering Wheel Angle Data Classification

According to the road video episodes, steering wheel angle data were classified into six categories: alert driving on straight road (AST), alert driving on curvature road (ACU), alert driving on connection road (ACO), fatigue driving on straight road (FST), fatigue driving on curvature road (FCU), and fatigue driving on connection road (FCO).

Results

According to the classification in Sect. 3.3, 12 sets of steering wheel angle data were selected to analyze the changes of steering wheel angle based on each category and the time interval of each group was 10s. The statistic analysis parameters, which include mean steering wheel angle, standard deviation of steering wheel angle and coefficient of variation of steering wheel angle, are listed in Table 1. ANOVA and paired sample t-tests were used to analyze the effects of driver state and road curvature on steering performance. Tukey's post hoc test was adopted to compare means for different factors. The significance level was adjusted for the number of comparisons being performed to preserve an alpha level of 0.05.

Mean Steering Wheel Angle

Effects of Road Curvature

The effects of road curvature on mean steering wheel angle were analyzed in different driving states. When driver was in alert state, paired sample t-tests were

Table 1 Statistical analysis result of steering angle

Data number	Mean						Standard deviation						Coefficient of variation			
	AST	ACU	ACO	FST	FCU	FCO	AST	ACU	ACO	FST	FCU	FCO	ACU	ACO	FCU	FCO
1	−1.104	−9.131	−5.371	−1.073	−12.399	−5.311	1.500	0.864	4.821	2.377	2.592	7.212	0.095	0.898	0.209	1.358
2	−0.447	−9.740	−5.969	−1.115	−11.308	−6.532	0.595	1.583	4.054	2.232	5.178	4.885	0.163	0.679	0.458	0.748
3	−1.083	−6.433	−8.252	−0.988	−12.687	−9.184	1.892	0.160	4.484	2.922	6.924	14.287	0.025	0.543	0.546	1.556
4	−1.284	−7.035	−9.314	−2.089	−9.852	−6.493	1.270	0.714	6.653	2.483	2.155	5.707	0.102	0.714	0.219	0.879
5	−1.636	−11.604	−11.330	−1.548	−13.043	−7.327	2.900	5.193	8.518	1.877	0.910	7.121	0.448	0.752	0.070	0.972
6	−1.924	−19.138	−10.197	−1.766	−11.456	−9.427	2.527	7.676	11.532	3.224	4.830	3.442	0.401	1.131	0.422	0.365
7	−1.171	4.836	−2.934	−1.653	−11.502	−7.911	1.849	2.090	2.656	2.845	4.447	6.804	0.432	0.905	0.387	0.860
8	−1.720	−6.641	−3.799	−2.540	−11.379	−8.797	3.316	2.176	5.364	3.277	3.299	6.736	0.328	1.412	0.290	0.766
9	−0.939	−7.193	−8.150	−2.814	−8.783	−4.829	2.029	3.385	4.311	4.066	6.271	6.458	0.471	0.529	0.714	1.337
10	0.848	−5.723	−5.568	−2.663	−6.746	−9.617	2.527	1.286	4.643	4.809	8.024	16.649	0.225	0.834	1.189	1.731
11	1.062	−9.515	−5.663	−4.063	−11.639	−10.274	1.832	4.845	6.203	1.858	1.069	6.169	0.509	1.095	0.092	0.600
12	−1.182	−10.281	−7.003	−2.336	−13.521	−8.097	2.445	4.449	10.521	6.026	5.489	7.546	0.433	1.502	0.406	0.932

used to compare means for the three types of road curvature based on the AST, ACO, and ACU data. Significant differences were found between the straight road and both the curvature road ($t(11) = 4.698$, $p = 0.001$) and the connection road ($t(11) = 8.803$, $p < 0.001$), and there was no significant difference between the curvature road and the connection road ($t(11) = -0.962$, $p = 0.357$).

When driver was in fatigue state, the analysis was based on the FST, FCO, and FCU data. Paired sample t-tests revealed that there were significant differences between the straight road and both the curvature road ($t(11) = 13.226$, $p < 0.001$) and the connection road ($t(11) = 12.095$, $p < 0.001$), and there was also significant differences between the curvature road and the connection road ($t(11) = -4.594$, $p = 0.001$).

Thus, mean steering wheel angle was affected by road curvature and there were significance differences between different road curvatures, no matter what driving state was.

Effects of Driver Fatigue

The effects of driving state on mean steering wheel angle were analyzed on different road curvatures. On the straight road, the mean of steering angle was lower when driver was in alert state than in fatigue state ($F(1,22) = 9.791$, $p = 0.005$), and there was no significant difference on the curvature road ($F(1,22) = 3.388$, p = 0.079) and on the connection road ($F(1,22) = 0.916$, $p = 0.349$).

In sum, mean steering wheel angle was affected by driving state only on straight road. When driver was fatigue, mean steering wheel angle was higher than alert driving, while there was no significant difference between the curvature and connection road.

Standard Deviation of Steering Angle

Effects of Road Curvature

The effects of road curvature on standard deviation of steering angle were analyzed in different driving states. When driver was alert, paired sample t-tests were used to compare means for the three types of road curvature. Significant differences were found between the connection road and both the curvature road ($t(11) = -6.458$, $p < 0.001$) and the straight road ($t(11) = -5.637$, $p < 0.001$), and there was no significant difference between the curvature road and the straight road ($t(11) = -1.369$, $p = 0.198$).

In fatigue driving state, a significant effect was found for the road curvature ($F(2,33) = 9.678$, $p < 0.001$). The Tukey post hoc test found that standard deviation of steering angle on the connection road was significantly different from both the straight road ($p = 0.001$) and the curvature road ($p = 0.008$), and there was no significant difference between the straight road and the curvature road ($p = 0.576$).

The effects on standard deviation of steering angle had significant differences between the connection road and both the curvature road and the straight road, while there was no significant difference between the curvature road and the straight road. It revealed that sharply fluctuated road curvature was the main influencing factor.

Effects of Driver Fatigue

The effects of driving state on standard deviation of steering angle were analyzed on different road curvatures. On the straight road, the standard deviation of steering angle of alert driving was significantly lower than that of fatigue driving ($F(1,22) = 6.981$, $p = 0.015$). There was no significant difference between the curvature road ($F(1,22) = 2.260$, $p = 0.147$) and the connection road ($F(1,22) = 1.410$, $p = 0.248$).

The standard deviation of steering angle was affected by driving state only on straight road when driver was fatigue, which was consistent with the effects of the mean steering wheel angle by driving state.

Coefficient of Variation of Steering Angle

As is discussed in Sects. 4.1 and 4.2, mean steering wheel angle and standard deviation of steering angle are affected by road curvature and driving state. Road curvature is the main factor, which lead to the change of steering wheel angle, changes in different driving state only on the straight road. In order to eliminate the effects of road curvature, coefficient of variation is presented to analyze the effect of driving state on the curvature road and the connection road. Coefficient of variation of steering angle is calculated according to CV_STEER = STD_STEER/MEAN_STEER, where CV_STEER represents coefficient of variation of steering angle, STD_STEER represents standard deviation of steering angle, MEAN_STEER represents mean steering wheel angle and the calculate results are listed in Table 1.

On the curvature road, there was no significant difference between alert driving and fatigue driving ($F(1,22) = 1.275$, $p = 0.271$). So was the connection road ($F(1,22) = 0.390$, $p = 0.239$).

Discussion

Driver fatigue affects driving performance seriously. The current research focuses on the correlation between steering related variables and driver fatigue, and the influence of road curvature is not considered. Mast et al. found that the lane-tracking ability decreased as the time on the task increased [14]. Variables

such as the times of lane departures, SDLP, and maximum lane deviation were found highly correlated with driver fatigue [23]. The mean square of lane deviation, mean square of high-pass lateral position, and SDLP showed good potential as drowsiness indicators [24]. However, Yang et al. [25] demonstrated that sleep deprivation had greater effect on rule-based than on skill-based cognitive functions: When drivers were sleep-deprived, their performance of responding to unexpected disturbances degraded, while they were robust enough to continue the routine driving tasks such as lane tracking, vehicle following, and lane changing [25].

Lane variability is widely considered to be an important metric of fatigue driving [26–28], but to date it required installation of lane-tracking cameras and complex video signal processing software. Video-based lane tracking is prone to data loss when lane markers are missing or covered (e.g., by sand or snow), when weather conditions are bad, or in darkness. These limit the application of lateral lane position variability for driver fatigue detection. Forsman et al. [1] found that steering wheel angle can be used effectively to estimate the relative changes in lateral lane position. Then, steering wheel angle was found to be the most sensitive indicators for driving fatigue detection [29–31]. In this study, the influence of driver fatigue and road curvature on steering wheel angle was analyzed.

This research developed a method to evaluate driver fatigue level, which is the precondition for extracting precise fatigue driving data. Previous works on fatigue level evaluation were primary based on psychomotor vigilance test (PVT) [32] and the KSS. In the literature of Otmani, subjects were invited to assess their level of alertness on the KSS every 10 min during the driving [8], which was more intrusive, and fatigue level and driving performance was affected by the assessment process. Another potential limitation is individual difference. Furthermore, Reyner and Horne [33] found that drivers sometimes underestimated the likelihood of falling asleep, despite feeling very sleepy. Forsman et al. [1] rated participant fatigue level by a 10-min PVT and KSS before and after the driving session, while, the driving session was only 30 min, it is not suitable for long driving. An objective evaluated method was proposed in our research, and fatigue level was scored based on KSS by three researchers through face video data during driving process of participants. In this method, driving fatigue level can be evaluated in real time without intrusion and individual difference.

Based on the KSS level and road curvature, driving simulator data were classified into six categories. On the same driving state, mean steering wheel angle revealed significant differences between different road curvatures. The analysis of standard deviation of steering angle on the alert driving state showed significant differences between connection road segments and both straight road and curved road segments. However, there was no significant difference between straight road and curved road segments. It suggests that steering wheel angle fluctuates stably in the same road curvature and the change of road geometry is the main cause of sharp fluctuation, which is consistent with Mortazavi [9].

On straight road segments, drivers showed significant differences in steering wheel angle, which include mean steering wheel angle and standard deviation of steering angle, during the fatigue driving state compared to alert driving. However,

there were no significant differences between curved road segments and connected road segments. In order to eliminate effects of road curvature, coefficient of variation of steering angle was proposed in our research. The experiment result is consistent with mean steering wheel angle and standard deviation of steering angle. It suggests that, fatigued drivers control ability deteriorates significantly on straight road segment, as they are more cautious and modify the steering wheel frequently on the curved road. Our research suggests that steering wheel variability could not be used directly for driver fatigue detection without considering road curvature.

Driver fatigue is the main factor of road crashes, consequently increasing researches focus on it. However, as applied psychological constructs, there is no generally accepted definition of driver fatigue. In our study, if KSS level is more than 7, driver is considered in fatigued driving state, and, if KSS level is less than 3, driver is considered in alert driving state. We are still not able to give an accurate definition, and the fatigue level is not determined when warn should be triggered as it is too dangerous to continue driving. The steering wheel angle data, which were used in our research, were collected from a driving simulator with realistic hardware. It is critical that a field study of real-world driving be conducted to assess the validity.

References

1. Forsman, P.M., B.J. Vila, R.A. Short, C.G. Mott, and H.P.A. Van Dongen. 2013. Efficient driver drowsiness detection at moderate levels of drowsiness. *Accident Analysis and Prevention* 50: 341–350.
2. Klauer, S.G., Dingus T.A., Neale, V.L., Sudweeks, J.D., and D.J. Ramsey. 2006. *The impact of driver inattention on near-crash/crash risk: An analysis using the 100-car naturalistic driving study data*. National Highway Traffic Safety Administration.
3. Coetzer, R.C., and G.P. Hancke. 2009. Driver fatigue detection: A survey. In *IEEE AFRICON Conference in AFRICON*, 1–6.
4. Nordbakke, S., and F. Sagberg. 2007. Sleepy at the wheel: Knowledge, symptoms and behaviour among car drivers. *Transportation Research Part F* 10 (1): 1–10.
5. Dinges, D.F., Mallis, M.M., Maislin,G., and J.W. Powell. 1998. Final report: Evaluation of techniques for ocular measurement as an index of fatigue and as the basis for alertness. Management. U.S. Department of Transportation, National Highway Traffic Safety Administration, Final Report # DOT HS 808 762, Washington, DC.
6. Dong, Y., Z. Hu, K. Uchimura, and N. Murayama. 2011. Driver inattention monitoring system for intelligent vehicles: A review. *IEEE Transactions on Intelligent Transportation Systems* 12 (2): 596–614.
7. Jin, L.S., Q.N. Niu, H.J. Hou, H.C. Xian, Y.L. Wang, and D.D. Shi. 2012. Driver cognitive distraction detection using driving performance measures. *Discrete Dynamics in Nature and Society* 2012: 1–12.

8. Otmani, S., T. Pebayle, J. Roge, and A. Muzet. 2005. Effect of driving duration and partial sleep deprivation on subsequent alertness and performance of car drivers. *Physiology & Behavior* 84 (5): 715–724.

9. Mortazavi, A., A. Eskandarian, and R.A. Sayed. 2009. Effect of drowsiness on driving performance variables of commercial vehicle drivers. *International Journal of Automotive Technology* 10 (3): 391–404.

10. Zhong, Y.J., Du, L.P., Zhang, K., and X.H. Sun. 2007. Localized energy study for analyzing driver fatigue state based on wavelet analysis. In *Proceedings of the 2007 International Conference on Wavelet Analysis and Pattern Recognition*, Beijing, 1843–1846.

11. Sayed, R., and A. Eskandarian A. 2001. Unobtrusive drowsiness detection by neural network learning of driver steering. In *Proceedings of the Institution of Mechanical Engineers, Part D: Journal of Automobile Engineering* 215(9), 969–975.

12. Sayed, R., Eskandarian, A., and M. Oskard. 2001. Driver drowsiness detection using artificial neural networks. In *Transportation Research Board 80th Annual Meeting*, Washington D.C.

13. Hulbert, S. 1972. Effects of driver fatigue. In *Human Factors in Highway Safety Research*, ed. T.W. Forbes, 110–132. New York: Wiley.

14. Mast, T.M., H.V. Jones, and N.W. Heimstra. 1966. Effects of fatigue on performance in a driving device. *Highway Research Record* 122: 93.

15. Kahneman, D. 1973. *Attention and Effort*. New Jersey: Prentice-Hall.

16. Wierwille, W.W., and L.A. Ellsworth. 1994. Evaluation of driver drowsiness by trained raters. *Accident Analysis and Prevention* 26 (5): 571–581.

17. Dingus, T.A., Hardee, L., and W.W. Wierwille. 1985. *Development of impaired driver detection measures*. USA Departmental Report 8504. Blacksburg: Department of Industrial Engineering and Operations Research, Virginia Polytechnic Institute and State University.

18. Elling, M., and P. Sherman. 1994. Evaluation of steering wheel measures for drowsy drivers. In *Proceedings of the 27th ISATA*, Aachen, Germany, 207–214. Germany: Automotive Automation Ltd.

19. Weller, G., and B. Schlag. 2009. A robust method to detect driver distraction. In *Proceedings Conference Human Centred Design for Intelligent Transport Systems*, 279–288.

20. Åkerstedt, T., and M. Gillberg. 1990. Subjective and objective sleepiness in the active individual. *International Journal of Neuroscience* 52: 29–37.

21. Nodine, E. 2008. The detection of drowsy drivers through driving performance indicators. Master's Thesis, Tufts University, United States.

22. Ingre M, Akerstedt T, Peters B. A. Anund, and G. Kecklund. Subjective sleepiness, simulated driving performance and blink duration: Examining individual differences [J], Journal of Sleep Research, 2006, 15(1): 47-53.

23. Yeo, M.V.M., X. Li, K. Shen, and E.P.V. Wilder-Smith. 2009. Can SVM be used for automatic EEG detection of drowsiness during car driving? *Safety Science* 47 (1): 115–124.

24. Stein, A.C. 1995. Detecting fatigued drivers with vehicle simulators. In *Driver impairment, driver fatigue and driving simulation*, ed. L. Hartley, 133–148. London: Taylor & Francis.

25. Yang, J.H., Z.H. Mao, L. Tijerina, T. Pilutti, J.F. Coughlin, and E. Feron. 2009. Detection of driver fatigue caused by sleep deprivation. *IEEE Transactions on Systems, Man and Cybernetics-Part A: Systems and Humans* 39 (4): 694–705.

26. Åkerstedt, T., M. Ingre, G. Kecklund, A. Anund, D. Sandberg, M. Wahde, P. Philip, and P. Kronberg. 2010. Reaction of sleepiness indicators to partial sleep deprivation, time of day and time on task in a driving simulator-the DROWSI project. *Journal of Sleep Research* 19 (2): 298–309.

27. Anund, A., G. Kecklund, B. Peters, Å. Forsman, A. Lowden, and T. Åkerstedt. 2008. Driver impairment at night and its relation to physiological sleepiness. *Scandinavian Journal of Work, Environment & Health* 34 (2): 142–150.

28. Sandberg, D., A. Anund, C. Fors, G. Kecklund, J.G. Karlsson, M. Wahde, and T. Åkerstedt. 2011. The characteristics of sleepiness during real driving at night-a study of driving performance. Physiology and subjective experience. *Sleep* 34 (10): 1317–1325.

29. Berglund, J. 2007. In-vehicle prediction of truck driver sleepiness-steering related variables. Master's Thesis, Linköping University, Department of Electrical Engineering, Sweden.
30. Friedrichs F., and B. Yang. 2010. Drowsiness monitoring by steering and lane data based features under real driving conditions, In *18th European Signal Processing Conference (EUSIPCO-2010)*, Aalborg, Denmark, 209–213.
31. Mattsson K. 2007. In-vehicle prediction of truck driver sleepiness. Master's Thesis, Luleå University of Technology, Sweden.
32. Dinges, D.F., and J.W. Powell. 1985. Microcomputer analyses of performance on a portable, simple visual RT task during sustained operations. *Behavior Research Methods, Instruments & Computers* 17: 652–655.
33. Reyner, L.A., and J.A. Horne. 1998. Falling asleep whilst driving: Are drivers aware of prior sleepiness? *International Journal of Legal Medicine* 111: 120–123.

Requirements Analysis of Operation Monitoring for Electric Vehicle in Transportation Industry

Cong Zhu, Lin Wang, Wengfeng Liu and Bin Li

Abstract Ensuring the operation efficiency and safety of electric vehicle is one of the main tasks of promoting the application of electric vehicle in transportation industry. Based on the operation characteristics of electric vehicle in transportation industry, the requirements of operation monitoring for electric vehicle were analyzed in this paper. Firstly, the operation processes of electric bus under typical urban conditions were simulated by the software of Advisor, and the potential faults of the battery pack during vehicle operation were demonstrated. Secondly, the peak power capability and driving range of electric bus under different conditions were studied by bench test and simulation software, and the effects of the state of charge (SOC) and temperature of battery pack on bus's operation capability were analyzed. Lastly, the aging characteristics of battery were researched by accelerated aging test, and the impacts of the battery aging on its ohmic resistance and available capacity were analyzed. The analysis results indicate that the change amplitude and frequency of the charging/discharging current of battery are drastic under typical urban driving conditions, which can easily lead to the battery risks such as over charge, over-discharge, and thermal runaway; both of the SOC and temperature of battery have obvious effect on the peak power capability and driving range of electric bus, and these two parameters can decline by about 50% when the battery temperature decreased from 20 to 0 °C; the loss of available capacity is mainly caused by the increase of ohmic resistance when the battery is aged, and the average loss rate of available capacity can be about 30–50% after one year of vehicle operation.

Keywords Electric vehicle · Transportation industry · Operation monitoring · Peak power · Driving range

C. Zhu · L. Wang (✉) · W. Liu · B. Li
Research Institute of Highway, Ministry of Transport, Beijing 100088, China
e-mail: wanglin@itsc.cn

Introduction

Now, the world is facing critical challenges of the oil depletion and environment pollution, and energy conservation and emission reduction are the common goals of the world automotive industry. Regarding to the Chinese automotive industry, promoting the development of new energy vehicle is especially urgent under the double pressures of air pollution and adjustment of automotive industry structure, and electric vehicle becomes one of the best breakthroughs for the development of new energy vehicle in China [1]. Since the Chinese government started the major special project of electric vehicle in 2001, the automotive enterprises, research institutes, and universities have enthusiastically studied and explored the key technologies in the field of electric vehicle. So far, the electric vehicle industry of China has made great achievements and has obtained more than 3000 patents of the key technologies of electric vehicle [2]. In some cities and areas, the application of electric vehicle in transportation industry has reached a certain scale [3], especially for the application in the urban bus field. At the end of 2014, there are 27 provinces in China that has applied electric bus to the transportation industry, and the number of city having the application exceeds 110, involving more than 300 bus enterprises with a total electric bus number of 36,617 and electric taxi number of 4000 [4]. According to the statistics of Ministry of Public Security, the new registered number of electric bus reached 92.5 thousand in 2015 [5]. According to the implement suggestion of promoting the application of new energy vehicle in transportation industry published by the Ministry of Transport, the number of new energy vehicle applied to transportation industry needs to reach 300,000. In summary, the application of electric vehicle in transportation industry has become an inevitable trend. As a new kind of carrying tool in transportation industry, ensuring the operation efficiency and safety is one of the main tasks in promoting the application of electric vehicle in transportation industry. Through remote monitoring the operation status of electric vehicle, the problems and faults during vehicle operation can be monitored and diagnosed in real time, and the related early warnings and remote services can be provided for the electric vehicle. This is very useful for ensuring the operation efficiency and safety of vehicle and has great significance for facilitating the application of electric vehicle in the transportation industry of China [6, 7].

There are many vehicle manufacturers at home and abroad that have successfully developed remote monitoring system for their vehicle products [8], such as the OnStar system developed by the General Motor of USA; the iDrive system developed by Bayerische Motoren Werke (BMW) of German can provide communication, positioning, navigation, and information services, and the OnStar system also can provide emergency rescue service; the G-Book system developed by Toyota of Japan can offer rescue, anti-theft tracking, and maintenance notice services; the Carwings system developed by Nissan of Japan can provide information consultant, environmental expert, and ternary security services; the InkaNet system developed by SAIC Motor of China can offer navigation, interphone, and

information services; the Incall system developed by Changan Motor of China can provide dynamic navigation, warning, tracking, and remote rescue services. However, all the monitoring systems mentioned above are applied to the conventional gasoline vehicles, and there are few such systems that are applied to the electric vehicles. Based on the operation characteristics of electric vehicles in transportation industry, the relationships between operation status and operation capability of electric vehicle are analyzed in this paper, to demonstrate the main requirements of operation monitoring and supply theoretical basis and guideline for the construction of the operation monitoring system for electric vehicle.

Charging/Discharging Characteristics

In order to reasonably test the performances of power, economy, and emission of vehicle, it has formed a standard driving condition set for the tests in transportation industry. The Urban Dynamometer Driving Schedule (UDDS), New York City Cycle (NYCC), and Extra Urban Driving Cycle of ECE (ECE-EUDC) are chosen from the standard set in this paper, to study the charging/discharging characteristics of electric vehicle in transportation industry. The methods for testing electric vehicle include real vehicle test, bench test, and software simulation. Generally, the economic and time costs of vehicle test and bench test are very high, while the costs of simulation are comparatively less with a reasonable accuracy. At present, there are many effective and reliable simulation softwares of electric vehicle existing in market. By the software of Advisor based on the MATLAB/Simulink, the operation processes of electric bus under typical urban driving conditions are simulated in this paper. The dynamic changes of charging/discharging current of battery pack are shown in Fig. 1, when the electric bus is operated under the three driving conditions mentioned above. The positive value of the vertical coordinates denotes the discharging current, while the negative value denotes the charging current in the figure. As shown in Fig. 1, the change amplitude and frequency of the charging/discharging current are drastic, and the maximum change amplitude can reach about 400 A. The main reason leading to the drastic change of current is that the electric bus needs to start/stop, accelerate, and brake frequently, which could result in a drastic change of the required power for the battery pack. This situation is especially obvious, when there is traffic congestion often occurred in the urban traffic system.

With the limits of the characteristics and electrochemical stability of battery electrode, there exists a limited value for the acceptable charging/discharging current of battery. The drastic change of current of electric bus operating under urban driving conditions can easily lead to overcharge or over-discharge of battery. According to the existing research results, the overcharge of battery can result in internal short circuit fault [9] and a rapid increase of internal temperature, even the risks of fire and explore [10]; the over-discharge of battery can decrease the available capacity and charge/discharge efficiency [11], increase the contact
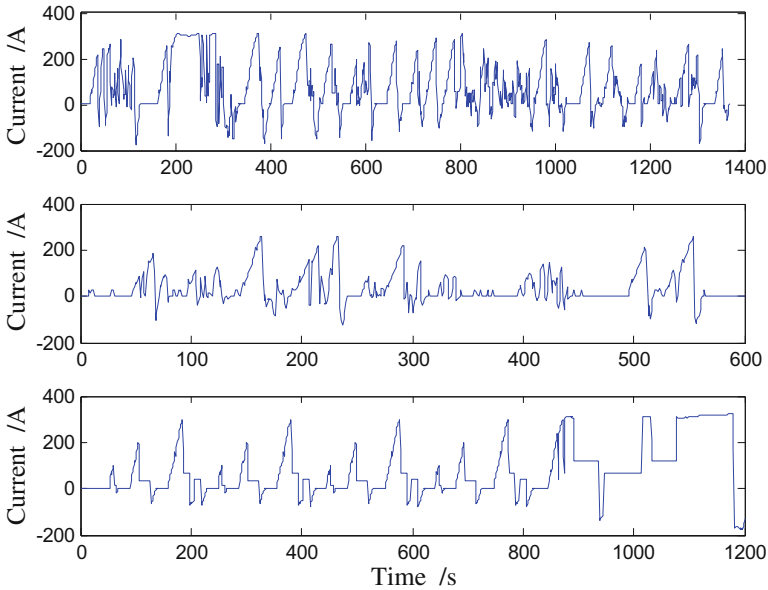
Fig. 1 Dynamic change of charging/discharging current under typical driving conditions

resistance of the positive electrode materials, and further reduce the life cycle of battery [9]. Although the electric bus is equipped with customized battery management system (BMS) to prevent overcharge and over-discharge of battery, the practical operation results of electric bus indicate that there is still overcharge or over-discharge occurred occasionally, due to the complicated chemical reaction inside the battery and the faultiness of BMS. Therefore, different from the conventional bus powered by fossil fuel, electric bus applied to transportation industry should be monitored to remedy the faultiness of BMS, for avoiding the occurrence of overcharge, over-discharge, and thermal runaway of battery during operation to the great extent.

Impact of Battery Temperature on the Operation Capability

When the electric vehicle is assigned for a given transportation task, the peak power capability and driving range of vehicle can be used as the main evaluation index of operation capability of the electric vehicle. The peak power capability can be used to evaluate that whether the electric vehicle can meet the power capability required by the operation task, and the driving range can be used to evaluate that whether the electric vehicle can meet the range requirement of the operation task.

At present, most researchers adopt the HPPC pulse test method given in the power battery test manual published by the Ministry of Energy of America, to test the peak power capability of battery of electric vehicle [12]. This method is used to test the power capability of an electric bus under different operating temperatures, and the test results are shown in Fig. 2. In the figure, T_{bat} denotes the battery temperature; P_{cmax} and P_{dmax} denote the peak charging and discharging power separately. As shown in Fig. 2, with the increase of battery SOC, the peak charging power decreases and the peak discharging power increases gradually; with the increase of battery temperature, both the peak charging and discharging power are increased. Using the P_{cmax} and P_{dmax} with the SOC of 50% as the basis, the SOC change of 10% can lead to a maximum change of about 20% of the P_{cmax} and P_{dmax}; when the T_{bat} decreases from 20 to 0 °C, the change amplitude of P_{cmax} and P_{dmax} can reach about 50%. Generally, the driving range of an electric bus is about 300 km, and the change rate of battery SOC is about 0.3%/km, while the change rate of battery temperature can be 5 °C/km. This shows the impact of battery temperature on the peak power capability of electric bus is comparable to the one of battery SOC.

The test methods of driving range of electric vehicle include constant velocity method and comprehensive condition method, and both the real vehicle test and software simulation adopt these two methods to study the driving range generally. In order to reasonably describe the driving range of electric vehicle in transportation industry, the America UDDS, Europe EUDC, and Japan 1015 driving conditions are chosen in this paper to study the driving range. The driving ranges of an electric bus under the three conditions mentioned above are calculated by the software of Advisor. The calculation results with battery temperature of 0, 20, and 40 °C are given in Table 1, and S in the table denotes the driving range. It is assumed that the initial battery SOC is 100%, and the battery temperature is constant during the
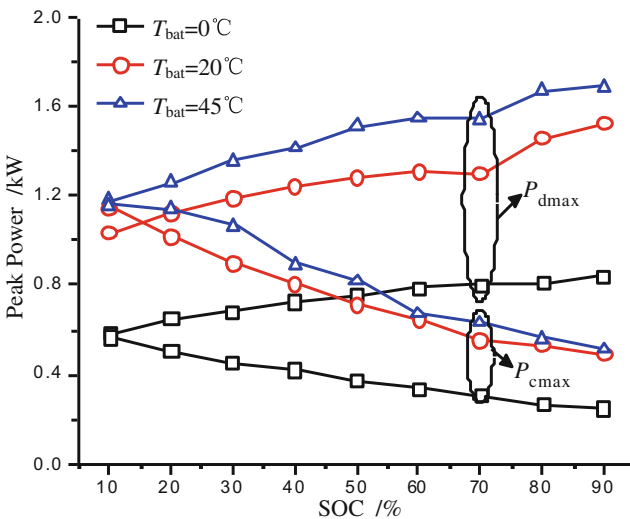
Fig. 2 Change of peak power capability with battery SOC and temperature

Table 1 Simulation results of the driving range of bus under different operating conditions

Driving condition T_{bat} (°C)	UDDS (km)	EUDC (km)	1015 (km)
0	$S = 52.5$	$S = 61.1$	$S = 53.5$
20	$S = 106.6$	$S = 118.8$	$S = 109$
40	$S = 122.9$	$S = 136.4$	$S = 125.8$

calculation. As shown in the table, the impact of driving condition on the driving range is relative small, and the change of S with the driving condition with the same temperature is around 15%; while the effect of battery temperature on the driving range is obvious. Using the driving range with battery temperature of 20 °C as the basis, when the battery temperature increases from 20 to 40 °C the driving range of electric bus increase 15.3, 14.8 and 15.4% separately; when the battery temperature decreases from 20 to 0 °C the driving range of electric bus decrease 50.8, 48.6 and 50.9% separately. This is because the change of battery temperature has a very obvious effect on the activity of electrode material. Low battery temperature leads to low activity of electrode activity and the available capacity of battery, and further results in an obvious decrease of the driving range of bus.

Generally, the vehicle managers estimate the operation capability of electric vehicles according to the battery SOC and history experience and establish the transport and charging schedules for the vehicle based on the estimation. For example, there is a transport task requiring a driving range of S_0, when the battery SOC is higher than a given low limited value of SOC_L, then the driving range is estimated to be higher than S_0, and the vehicle is estimated to be capable of completing the transport task. However, according to the above analysis, the operating capability of electric vehicle is not only affected by the battery SOC, but also affected by the battery temperature. Estimation of operation capability of electric vehicle by history experience requires high demands on the professional knowledge of vehicle manager. Through monitoring the battery SOC and temperature during the operation of electric vehicle, and with the combination of the knowledge base of peak power capability and driving range, the operation capability of electric vehicle can be reasonably estimated by the managers. Reasonable estimation is helpful for the dispatch and management of electric vehicle in transportation industry and useful for taking full advantage of electric vehicle in saving energy and reducing emission, and further enhancing the promotion of the application of electric vehicle in transportation industry.

Impact of Battery Aging on the Operation Safety

The practical operation process of electric vehicle in transportation industry indicates that with the increase of operation time, the vehicle battery will inevitably become aged. When the battery is aged, the internal resistance of battery is

increased, and the available capacity of battery is decreased gradually. In order to reasonably describe the dynamic change of the battery parameters during the process of battery aging, the aging characteristics of battery are studied by bench test of accelerated aging. The changes of the ohmic resistance and available capacity of battery with the operation time of electric vehicle are shown in Figs. 3 and 4 [13]. As shown in the figures, the ohmic resistance and the average loss rate of capacity are increased with the increase of vehicle operation time; generally, the increment of battery ohmic resistance is about 20% when the electric vehicle operates for one month, while the average loss rate of battery capacity can be around 30–50%, when the electric vehicle operates for one year.

Fig. 3 Variation of ohmic resistance of battery with the operation time
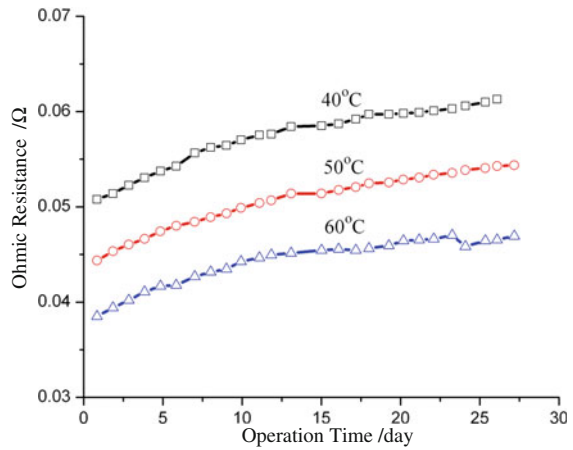
Fig. 4 Variation of average rate of capacity loss with the operation time
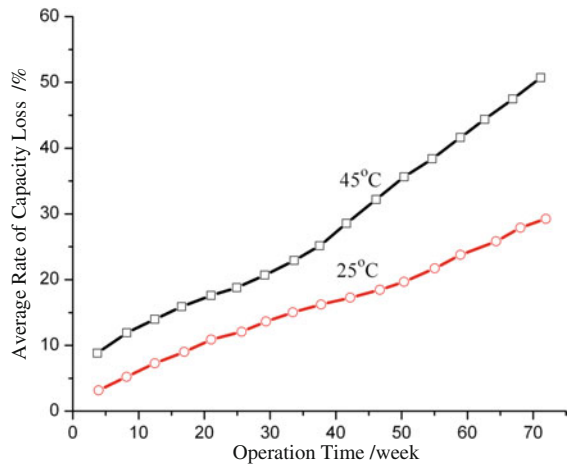
The ohmic resistance and available capacity of battery are the main basis for BMS to provide energy and thermal management. The manufacturers of electric vehicle generally establish corresponding strategies for the energy and thermal management according to these two battery parameters. The process of establishing strategies is called as the matching and calibration of BMS. The quality of matching and calibration has an important impact on the operation performance of electric vehicle, and especially on the operation safety of vehicle. As analyzed above, the parameters mentioned above show a obviously nonlinear change during the process of battery aging. Although there are many prediction models are developed to adapt the change of the parameters for BMS, and to support the online calibration of BMS, the practical application of the models on electric vehicle is hard to be realized due to the low accuracy and reliability of the models. If the vehicles are sent back to the manufactures for recalibration, it must lead to huge economic and time costs. Through monitoring the electric vehicle in transportation industry, the ohmic resistance and available capacity of battery can be reasonably estimated based on the large amount of history operation data, to support the online calibration of BMS and ensure the operation safety of electric vehicle in transpotation industry.

Conclusion

When the electric vehicles of transportation industry operate under the typical urban driving conditions, the drastic change of power requirement for vehicle will lead to a drastic change of charging/discharging current. This situation can easily result in the overcharge and over-discharge of battery. Through monitoring the charging/discharging current during vehicle operation, the faults of overcharge, over-discharge, and thermal runaway can be avoided as much as possible.

Both the battery SOC and temperature have complicated and obvious impact on the operation capability of electric vehicle. Through monitoring the dynamic change of battery SOC and temperature during vehicle operation, the operation capability of vehicle can be reasonably estimated with the knowledge base of the relationship between battery SOC, temperature, and operation capability. Then, the transport task can be reasonably assigned for the electric vehicle, and the advantages of vehicle can be fully taken.

With the increase of operation time of electric vehicle, the internal resistance of battery will increase and the available capacity of battery will decrease gradually. This has adverse effect on the reliability of BMS. Through monitoring the aging degree of battery during vehicle operation, the online calibration of BMS can be effectively completed accordingly. This is helpful for providing reasonable strategies of energy and thermal management by BMS and ensuring the operation safety of electric vehicle in transportation industry.

References

1. Dongpeng, Yue, Hao, Zhiyong, and Zhang, Junzhi. 2004. Hybrid-electric vehicle research development and tendency. *Tractor & Farm Transporter* 7: 1–4.
2. Hailong, Jiang., and Wei, Ruibin. 2013. Statistical analysis of electric vehicle patents in China. *Modern Information* 33 (3): 168–172.
3. Qian, Bin., Shi, Dongyuan., and Xie, Pingping., et al. 2014. Optimal planning of battery charging and exchange stations for electric vehicles. *Automation of Electric Power Systems* 38 (2): 64–69 + 84.
4. Xu, Chengguang. At the end of the total number of electric bus in China reaches 60, 000. [EB/OL]. http://www.d1ev.com/41954.html, 2016-01-21.
5. China Electric Vehicle Net. The registered number of electric bus in China is about 92, 500 in 2015. [EB/OL]. http://news.ddc.net.cn/newsview_66199.html, 2016-02-29.
6. Xie, Hui., Xiao, Bin., and Hao, Mingde., et al. 2006. Development of a wireless remote monitoring system for electric vehicle demonstration operation. *Automotive Engineering* 28 (8): 734–737 + 770.
7. Jin, Li. 2011. Research on matching relations between operation and battery state of electric bus. Master Dissertation of Beijing Jiaotong University, 2011.
8. Xiong, Yan. 2012. Research on the impact of mobile communications technology for automotive information. Master Dissertation of Jilin University.
9. Zhenping, Zhou., Zhao, Shixi., Liu, Zhen., et al. 2001. Study on mechanism of capacity loss for spinel $Li_xMn_2O_4$. *Materials Review* 15 (5): 30–33.
10. Hongyu, Chen, Tang, Zhiyuan., Lu, Xinghe, et al. 2006. Research of explosion mechanism of lithium-ion battery. *Progress in Chemistry* 18 (6): 823–831.
11. Yu, Zhongbao., Hu, Junwei., Chu, Xuguang., et al. 2006. Effects of over-discharge on performance of MCMB-$LiCoO_2$ lithium-ion battery. *Chinese Battery Industry* 11 (4): 223–226.
12. Idaho National Laboratory. 2010. Battery test manual for plug-in hybrid electric vehicles. http://www.inl.gov.
13. Christophersen, J., Bloom, I., and Thomas, E, et al. 2010. Developing modeling capability to predict life of batteries. US-China Electric Vehicle and Battery Technology Workshop, Argonne National Laboratory, USA.

Study on the Establishment of Vulnerability Source Evaluation Model for the Road Traffic Safety

Tingting Gao, Wuhong Wang and Min Li

Abstract The traffic system is an extremely complex coupling system. This chapter establishes the traffic system vulnerability model based on the complexity of the system and analyzes the vulnerability source. The vulnerability source model of traffic safety is set up by employing fuzzy analytical hierarchy process. The study establishes secondary vulnerability source indexes, which consist of human, vehicle, road and environment, and third vulnerability source indexes, which consist of driver and non-driver (human); motorized vehicle and non-motorized vehicle (vehicle); urban road and highway (road); controllable environment and non-controllable environment (environment). Conclusions demonstrate that the influence degree of the road traffic safety's vulnerability source is from the big to the small: the driver, the motorized vehicle, the highway, the non-driver, the non-motorized vehicle, the controllable environment, the urban road and the non-controllable environment. The result of the evaluation for the model is similar to the current situation of road traffic safety, and thus, it can be taken as a valuable reference to promoting the traffic safety.

Keywords Safety engineering · Traffic safety · Vulnerability source · FAHP · Order

Introduction

Road traffic system is a complex system of people, vehicles, roads and environment. Complex systems usually exhibit a variety of complex behaviors, but they have some common characteristics, that is, the mechanism underlying these phenomena is

T. Gao (✉) · W. Wang (✉) · M. Li
School of Mechanical Engineering, Beijing Institute of Technology,
Beijing 100081, China
e-mail: gaotingting@tute.edu.cn

W. Wang
e-mail: wangwuhong@bit.edu.cn

T. Gao
School of Auto and Transportation, Tianjin University of Technology and Education,
Tianjin 300222, China

instability [1]. According to the research of Hongzhang [2] on complex systems, the process of operation of the sudden collapse of the risk is known as brittle. For example, once the traffic congestion occurred due to traffic accidents, the transportation system temporarily loses its original function, when the fragility of the traffic system is excited [3–5]. Therefore, analyzing the brittle source of road traffic safety has great significance to improve the road traffic safety.

So far, domestic and foreign scholars on road traffic safety analysis and research used a variety of methods: according to the time of the accident is divided into pre- and post-analysis; in accordance with the scope of the study is divided into macro- and micro-analysis. The main methods used in China are fuzzy evaluation method, gray comprehensive evaluation method and extension evaluation method [6–8]. In foreign countries, DEA (data envelope analysis) and SEM (structure equation model) and so on [9, 10]. But the common ground of the above studies is that the influence index of road traffic safety is subdivided into more detailed indexes, so the model is established by combining the qualitative and quantitative analysis methods, especially for people, vehicles, roads and environment. There are no uniform sub-criteria. Therefore, based on the characteristics of previous studies, this chapter standardizes the index of the subdivision, but the number of indicators is less, and the influence factors are more comprehensive. The brittleness model of the complex system is used to analyze the brittle source of road traffic safety, and the index of fragile source is sorted by fuzzy analytic hierarchy process (AHP).

Transportation Complexity

Complexity Judgment

Road traffic system is a strong coupling system composed of human, vehicle, road and environment. It has obvious openness, complexity, evolutionary emergence, hierarchical and huge quantity. The complexity of the system is shown in formula (1) [11].

$$c = \frac{m}{n-1} \qquad (1)$$

where

c complexity;
m the number of actual relationships between elements within the system;
$n-1$ the smallest possible number of relations, the number of elements minus 1.

When

$c = 1$ the system is a simple system;
$c > 1$ the system is a complex system.
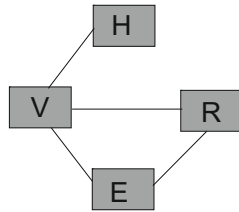
where: H—Human, which exist in the transportation;

V—Vehicle, which exist in the transportation;

R—Road, which exist in the transportation;

E—Environment.

$c = \dfrac{4}{3}$ is given by formula 1in the traffic system, so the traffic system is complex system.

Fig. 1 Traffic system

The drivers in the car of the road traffic system exist in the road and the environment; the relationship between them is shown in Fig. 1:

$c = \frac{4}{3}$ is given by Formula 1 in the traffic system, so the traffic system is a complex system.

Vulnerability Model of Traffic

The graph theory is an intuitionistic mathematical model which expresses the relations among the elements in the traffic system. Therefore, the vulnerability model of the traffic system is established by graph theory. The traffic system model is shown in Fig. 2a is for the original model diagram, in order to facilitate the calculation, the traffic system simplified model as shown in (b). The traffic system has four components, and the relationship between the various components is as follows:

Fig. 2 Traffic system model
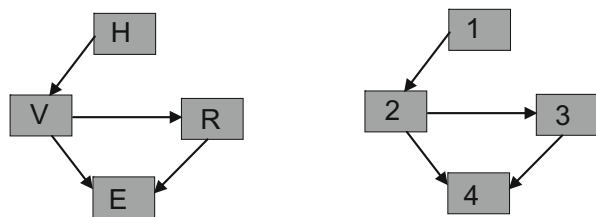
(a) Original model   (b) Simplified model

(1) The nodes of this system: $J = \{1, 2, 3, 4\}$
(2) The directional branch of the system: $Z = \{(1, 2), (2, 3), (2, 4), (3, 4)\}$

(3) Adjacency matrix of the system: $L = \begin{bmatrix} 0 & 1 & 0 & 0 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \\ 0 & 0 & 0 & 0 \end{bmatrix}$

(4) The reachable matrix of the system: $L + I = \begin{bmatrix} 1 & 1 & 0 & 0 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

$$(L+I)^2 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} \quad (L+I)^3 = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix} = (L+I)^2$$

So the reachable matrix of the system is: $K = \begin{bmatrix} 1 & 1 & 1 & 1 \\ 0 & 1 & 1 & 1 \\ 0 & 0 & 1 & 1 \\ 0 & 0 & 0 & 1 \end{bmatrix}$

In the road traffic system, human being is the main brittle source of the road traffic system. The collapse of this factor leads to the collapse of the vehicle, the road and the environment. At the same time, the collapse of the traffic system leads to the collapse of the traffic system. The road traffic system is a third brittle source, and its collapse will lead to the collapse of the road and the environment, road traffic system is the third brittle source of the collapse, he will lead to the collapse of the road traffic system is not the driver The collapse of the environment.

Evaluation Model of Vulnerability Source of Traffic System

The road traffic system has strong layering, and its factor, order of brittle source, is very fuzzy, so it can be used to evaluate the vulnerability of the traffic system. Specific steps are as follows:

The Establishment of Traffic System Vulnerability Source Level Map

Level evaluation can be generally divided into three layers: the highest layer for the evaluation of the target, the middle layer of factors affecting the target layer and the bottom layer of indicators for the impact of the middle layer, usually with the structure diagram.

Construction of the Priority Matrix

The relative importance of each factor to the next factor is expressed in matrix form, as shown in Table 1.

Fuzzy Consistent Matrix

$$R = \left[r_{ij}\right]_{m \times m} = \left[\frac{r_i - r_j}{2m} + 0.5\right]_{m \times m} \tag{2}$$

where

m	the number of rows of the matrix
r_i, r_j	the matrix is summed by rows, namely the sum of the i row or j rows, respectively r_i, r_j

The Relative Importance

The relative importance of the weight vector is calculated as follows:

The elements of the fuzzy matrix are multiplied by a new vector, and then, each component of the new vector is n squared. Finally, the resulting vector is normalized to the relative importance w_s:

Table 1 Selection of e_{ij} value

	e_{ij}
i factors are more important than j factors	1.0
j factors are more important than i factors	0
i factors and j factors are equally important	0.5
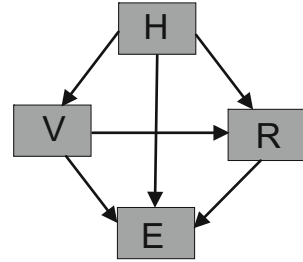
$$w_s = \frac{\left(\prod_{j=1}^{n} r_{ij}\right)^{1/n}}{\sum_{k=1}^{n}\left(\prod_{j=1}^{n} r_{kj}\right)^{1/n}} \tag{3}$$

Hierarchical Ordering

The relative importance of the elements of each index w_s multiplied by the corresponding index is the total ranking of the corresponding element to the target layer (Fig. 3).

The Vulnerability Source Evaluation of Road Traffic Safety

Road Traffic Accident Hierarchy Figure

The main factors leading to road traffic accidents were human, vehicles, roads and the environment. In order to comprehensively consider the influence of various factors of the road traffic safety and the structure of the priority matrix, the secondary indicators of human, vehicles, roads and the environment, which are the main factors that impact on the road traffic safety, are as follows: non-vehicle owners (including pedestrians, bicycles (also electric), tricycles, etc.) and vehicle owners; motor vehicles (including passenger cars, trucks, motorcycles, etc.) and non-motor vehicles; city roads (including the general urban roads and urban expressways) and highways (first class road, second class road); controllable environment (road lighting, traffic control devices and road protection facilities) and uncontrollable environment (weather conditions). Road traffic accident is divided into A, B and C: A is the highest layer of the road traffic accident; B is the middle layer and consists of $B1$ (human), $B2$ (vehicle), $B3$ (road) and $B4$ (environment); $B1$ includes $C1$ (driver) and $C2$ (non-driver), $B2$ includes $C3$ (motor vehicle) and $C4$ (non-motor vehicle), $B3$ includes $C5$ (city road) and $C6$ (highway), $B4$ includes $C7$ (controlled environment) and $C8$ (uncontrolled environment), which are shown in Fig. 4.

(a)

the cause of traffic accident

(b)

| human | vehicle | road | environment |

(c)

driver | non-driver | motor | non-motor | city road | highway | conttrolla ble | uncontroll able
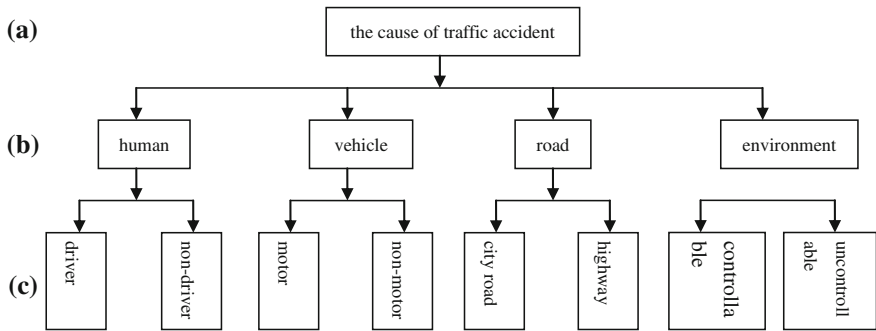
Fig. 4 Hierarchical structure of traffic crash

Model

Build the Priority Matrix

$$A - B = \begin{bmatrix} 0.5 & 1.0 & 1.0 & 1.0 \\ 0 & 0.5 & 1.0 & 1.0 \\ 0 & 0 & 0.5 & 1.0 \\ 0 & 0 & 0 & 0.5 \end{bmatrix}$$

$$B1 - C = \begin{bmatrix} 0.5 & 1.0 \\ 0 & 0.5 \end{bmatrix} \quad B2 - C = \begin{bmatrix} 0.5 & 1.0 \\ 0 & 0.5 \end{bmatrix}$$

$$B3 - C = \begin{bmatrix} 0.5 & 0 \\ 1.0 & 0.5 \end{bmatrix} \quad B4 - C = \begin{bmatrix} 0.5 & 1.0 \\ 0 & 0.5 \end{bmatrix}$$

Fuzzy Consistent Matrix

Fuzzy consistent matrix is calculated by formula (2)

$$E_{AB} = \begin{bmatrix} 0.5 & 0.625 & 0.75 & 0.875 \\ 0.375 & 0.5 & 0.625 & 0.75 \\ 0.25 & 0.375 & 0.5 & 0.625 \\ 0.125 & 0.25 & 0.375 & 0.5 \end{bmatrix}$$

$$E_{B_1C} = \begin{bmatrix} 0.5 & 0.75 \\ 0.25 & 0.5 \end{bmatrix} \quad E_{B_1C} = \begin{bmatrix} 0.5 & 0.75 \\ 0.25 & 0.5 \end{bmatrix}$$

$$E_{B_3C} = \begin{bmatrix} 0.5 & 0.25 \\ 0.75 & 0.5 \end{bmatrix} \quad E_{B_4C} = \begin{bmatrix} 0.5 & 0.75 \\ 0.25 & 0.5 \end{bmatrix}$$

Table 2 Relative proportions of traffic safety vulnerability source

First-grade indexes	Weight	Second-grade indexes	Weight	Total weight
Human	0.3527	Driver	0.634	0.2236
		Non-driver	0.366	0.1291
Vehicle	0.2854	Motor	0.634	0.1809
		Non-motor	0.366	0.1045
Road	0.2168	City road	0.366	0.0793
		Highway	0.634	0.1375
Environment	0.1415	Controllable environment	0.634	0.0920
		Uncontrollable environment	0.366	0.0531

Hierarchical Ordering

Hierarchical order is using formula (3)

$$w_{AB} = (0.3527, 0.2854, 0.2168, 0.1451)^{\mathrm{T}}$$

$$w_{B_1C} = (0.6340, 0.3660)^{\mathrm{T}} \quad w_{B_2C} = (0.6340, 0.3660)^{\mathrm{T}}$$

$$w_{B_3C} = (0.3660, 0.6340)^{\mathrm{T}} \quad w_{B_2C} = (0.6340, 0.3660)^{\mathrm{T}}$$

The final order is shown in the last column of Table 2.

Through the calculation and analysis, it is concluded that in the first-grade index of influencing the vulnerability source of road traffic safety, human (0.3527) > vehicle (0.2854) > road (0.2168) > environment (0.1415), which also verified the results of the vulnerability source model of road traffic system in Fig. 3; vulnerability source secondary index rankings of the road traffic safety are driver (0.2236)⟩ motor vehicle (0.1809)⟩ highway (0.1375)⟩ non-driver (0.1291)⟩ non-motor (0.1045)⟩ controllable environment (0.0920)⟩ city road (0.0793)⟩ uncontrollable environment (0.0531). This result is consistent with the current situation of road traffic safety.

Conclusion

(1) Based on the vulnerability model of complex system, the complexity of road traffic system is quantitatively defined, which is more convincing than qualitative description. The vulnerability source of the road traffic system is analyzed and calculated by graph theory, and the main vulnerability source is obtained, followed by vehicle, road and environment.

(2) Based on fuzzy AHP, the vulnerability source evaluation model of road traffic system is established and calculated. This method makes up for the shortcomings of AHP, which requires expert scoring and strong subjectivity; the calculation is simple; and the model of road traffic vulnerability is also verified.

References

1. Song, Wang, and Wang Ying. 2011. Causation analysis of complex system safety accident based on brittle structure collapse. *China Safety Science Journal* 21 (5): 138–142.
2. Hongzhang, Jin, and L. Wu Hongmei. 2005. Analyzing brittleness factors of complex systems. *Journal of Harbin Engineering University* 26 (6): 739–743.
3. Kunnrni, Yeo., and Chi-Hyuck, Jun. 2002. Teletraffic analysis of cellular communication systems with general mobility based on hyper-Erlang characterization. *Computers & Industrial Engineering* 42: 507–520.
4. Kudoh, Yuki, Hisashi Ishitani, and Ryuji Matsuhashi. 2001. Environmental evaluation of introducing electric vehicles using a dynamic traffic-flow. *Applied Energy* 69: 145–159.
5. Mihiya, Takahashi., Toshio, Nakanishi., Isao, Miyoshi., and Tomokazu, Fujikura. An evaluation of the road traffic system simulator PIMTRACS by PIM. *Mathematics and Computers in simulation* 59: 45–56.
6. LinYu, Niu Jianfeng, and Xu Ying. 2011. Projection pursuit method for road traffic safety macroscopic evaluation. *Journal of Safety and Environment* 11 (2): 221–223.
7. Rende, Yu., and Shi Peng. 2008. The research on road traffic security evaluation based on the theory of fuzzy. *Mathematics in Practice and Theory* 38 (7): 109–116.
8. Manrong, Yuan, Cheng Wei, and Zhao Junjun. 2013. Safety evaluation model of the urban road intersections based on extenics. *Mathematics in Practice and Theory* 43 (20): 26–31.
9. Hassan, Y., and G. Gihreel. 2000. Evaluation of highway consistency and safety practical application. *Journal of Transportation Engineering* 3: 127–131.
10. Al-Refaie, Abbas. 2013. Factors affect companies' safety performance in Jordan using structural equation modeling. *Safety Science* 57: 169–178.
11. Hongzhang, Jin, and WeiQi. 2010. *The Brittleness Theory and Application of Complexity*. Xian: Northwestern Polytechnical University Pres.

ABS Self-Adjustment Threshold Control Based on MATLAB

Neng Wan and Jian Xiong

Abstract In order to get good control effects on ABS simulation study, the optimal logic threshold value of logic threshold value control method will be changed with the variation in speed and road adhesion coefficient. In order to study on self-adjustment logic threshold control, the vehicle model with wheeled three degrees of freedom and effective logic threshold limiting value control strategy were created in MATLAB. Simulation tests were carried out in different vehicle speed and road adhesion coefficients; then, select two important parameters which are wheel slip rate and wheel speed to study, by analyzing the duration of optimal slip rate and the stability of wheel rotation, verified the effectiveness of self-adjustment logic threshold control.

Keywords ABS · Self-adjustment control · Logic threshold control

Introduction

With the development of ABS (anti-lock braking system), various control algorithms constantly emerges, such as logic threshold value control, PID control, slip mode control, and fuzzy control [1–4]. However, some of the algorithms are too complex and limited by theoretical research conditions, and the new control algorithm cannot be used by ABS products. At present, the logic threshold control algorithm is the most widely used algorithm in domestic and abroad, which is characterized by its low dependence on the establishment of the mathematical

N. Wan (✉)
College of Mechanics Engineering and Transportation, Southwest Forestry University, Kunming 650224, China
e-mail: charles244@163.com

J. Xiong
Faculty of Transportation Engineering, Kunming University of Science and Technology, Kunming 650500, China
e-mail: xjebox@163.com

model and the high efficiency of the control algorithm in the system. Using the logic threshold control algorithm, just use the vehicle wheel angular acceleration and the slip rate threshold value control in braking process—it can achieve the basic anti-lock braking cycle. In the majority of scholars research algorithm, the control algorithm and the logic threshold are usually presented for a single speed or road condition; however, in different speed and different adhesion coefficients of road, different threshold values are needed in order to meet the demands of effective control.

In this paper, a self-adjustment logic threshold control model is established; the simulation test is carried out under the conditions of different vehicle speed and road adhesion coefficients. The effectiveness of the self-adjustment threshold control is verified.

Establishing the Mathematical Model of Vehicle

In order to facilitate ABS modeling and not to lose its authenticity, this paper established the two-wheel-vehicle model; the advantage is ignoring the changes in the lateral force; at the same time, model using longitudinal Magic Formula can accurately describe tire braking force and motion of the single operating condition, which is suitable to process and analyze the braking performance of ABS and to simplify the problem. The force analysis of vehicle braking process is shown in Fig. 1.

Equation of vehicle motion:

$$\frac{\mathrm{d}vm}{\mathrm{d}t} = -F_\mathrm{f} - F_\mathrm{b} \tag{1}$$

where F_f is the front-wheel braking force, F_b the rear-wheel braking force, m the quality of the vehicle, v the speed, and t the braking time.

The breaking force is represented by the vertical load and the road adhesion coefficient:

$$\begin{cases} F_\mathrm{f} = F_\mathrm{zf}\psi_\mathrm{f} \\ F_\mathrm{b} = F_\mathrm{zb}\psi_\mathrm{b} \end{cases} \tag{2}$$

Fig. 1 Force analysis of braking
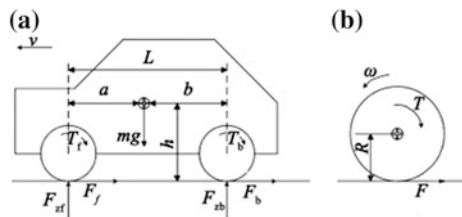
where F_{zf} and F_{zb} are the vertical load of the front and rear wheels of the vehicle, $F_{zf} = \frac{b}{L}mg + m\frac{dv}{dt}\frac{h}{L}$, $F_{zb} = \frac{a}{L}mg - m\frac{dv}{dt}\frac{h}{L}$; ψ_f and ψ_b are the adhesion coefficients of the front and rear wheels of the vehicle.

Simultaneously, Eqs. 1 and 2 are obtained for the differential equations of vehicle braking:

$$\frac{dv}{dt} = \frac{-g(\psi_f b - \psi_b a)}{L - h(\psi_f - \psi_b)} \tag{3}$$

a is the horizontal distance between the front axle and the barycenter of the vehicle, b is the horizontal distance between the rear axle and the barycenter of the vehicle, h is the vertical distance between the barycenter and the road, and L is the wheelbase.

The relationship of the road adhesion coefficient ψ and the wheel slip rate δ can be expressed as:

$$\psi = \psi_H e^{(n_1 \delta^{n_2})} - C \tag{4}$$

where ψ_H is the peak of the wheel adhesion coefficient under high speed; n_1 and n_2 are the coefficients; and C is the modification factor.

The relationship of the peak of the wheel adhesion coefficient and speed can be expressed as:

$$\psi_H = \psi_L - C_1 \tag{5}$$

where ψ_L is the peak of the wheel adhesion coefficient under low speed; C_1 is the constant.

The relationship of the modification factor C and the wheel slip rate δ can be expressed as:

$$C = \begin{cases} 0 & (s < 0.2) \\ (s - 0.2)\,C_2 & (s \geq 0.2) \end{cases} \tag{6}$$

In the formula, C_2 is the constant.

In the Fig. 1b, the wheel rotation equation is given below:

$$FR - T = \omega J \tag{7}$$

where F is the road breaking force; R the radius of the wheel; J the moment of inertia of the wheel; T the wheel braking torque; and ω the angular velocity of the wheel.

The formula of the wheel slip rate [5–6] is:

Fig. 2 Relationship curves of
ψ and δ
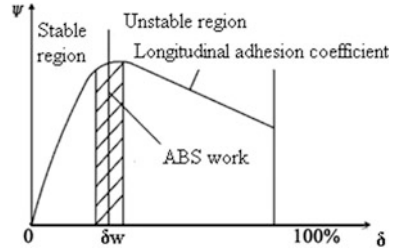
$$\delta = \frac{v - \omega R}{v} \qquad (8)$$

The relationship curves of the road adhesion coefficient ψ and the wheel slip rate δ are shown in Fig. 2. δ_w is the best control point of the slip rate, with larger longitudinal adhesion coefficient. In the braking, as far as possible to keep the slip rate δ in the vicinity of the δ_w. In the high adhesion coefficient road condition, δ_w is about 20%.

Using the "m" file in MATLAB software, establish a mathematical simulation model of the formulas 3–8 and perform a simulation analysis under different vehicle speed and adhesion coefficients, through wheel angular velocity curve to determine the wheel angular acceleration threshold value a_1 of optimal wheel slip rate and two boundary region wheel angular acceleration threshold values a_2 and a_3 of ABS working conditions. The wheel angle acceleration threshold value is stored in the above condition, and it is used as the reference threshold value under various conditions in the adaptive threshold control.

Due to the slip rate threshold δ_1 is determined based on the different road conditions, through the simulation experiment, under different conditions in the braking without ABS, the slip rate at which the longitudinal force reaches the peak time is defined as the optimum slip rate under this condition. From the high road adhesion coefficient to the low, the best slip rate is reduced from 20 to 15%; therefore, the best slip rate is also changed with the change in working condition.

The Establishment of ABS Logic Threshold Control and Adaptive Adjustment Algorithm

1. ABS logic threshold control

Select the ABS logic threshold control strategy and set the wheel angular velocity threshold value of a_1, a_2, and a_3 and reference slip rate threshold δ_1 of control system.

The control process is as follows: In the early stage of braking, the ABS does not work, and the brake pedal to generate the braking force directly output to the

system. When detecting the wheel angular acceleration $a < a_1$, the ABS starts to work and braking force will not output to the brake system, and the system enters the holding pressure stages, until $\delta > \delta_1$ into the reducing pressure stages; when $a > a_1$, braking force is again acting on the braking system. When $a < a_2$, or $\delta > \delta_1$, braking system starts reducing pressure, braking force will not output to the brake system; until $a > a_3$ or $\delta < \delta_1$, braking force is again acting on the braking system, then start the next cycle. In the circulation process, according to the road adhesion coefficient adjust threshold value at any time, ABS stops working when reduced to the lowest speed or suddenly removed the braking force.

2. Adaptive adjustment algorithm

Under different roads and speed, often require different threshold values [7], in order to achieve good control effect under different roads and speed; this paper proposes a threshold self-adjusting control, and the control can select and adjust the wheel's angular acceleration threshold value under the different roads and speed conditions [8]. Specific methods are as follows:

(1) Simulating through MATLAB software, obtain multiple groups of ABS threshold value under a variety of speed and road adhesion coefficient conditions; these conditions form the basic covering of the normal driving situations of vehicle.

(2) Analyzing the simulation diagram, the critical points of the single threshold value control failure of various conditions are used as the basis for dividing each range. The speed is divided into several speed range; the maximum speed difference of each speed range is 20 km/h. Each speed range, according to the road adhesion coefficient, is divided into three parts: low adhesion coefficient, medium adhesion coefficient, and high adhesion coefficient. The range distribution is shown in Table 1.

Table 1 Distribution range of adaptive adjustment algorithm

Speed range/ (km/h)	Road adhesion coefficient range	Threshold value			
		a_1/ (rad/s^2)	a_2/ (rad/s^2)	a_3/ (rad/s^2)	a_4/ (rad/s^2)
30–50	Low(0.0–0.3)	−23	−12	15	15
	Medium (0.3–0.6)	−35	−18	15	17
	High (0.6–1.0)	−45	−24	15	20
50–70	Low (0.0–0.3)	−25	−14	15	15
	Medium (0.3–0.6)	−27	−20	15	17
	High (0.6–1.0)	−50	−27	15	20
70–90	Low (0.0–0.3)	−30	−15	15	15
	Medium (0.3–0.6)	−45	−22	15	17
	High (0.6–1.0)	−61	−30	15	20
>90	Low (0.0–0.3)	−35	−23	15	15
	Medium (0.3–0.6)	−52	−34	15	17
	High (0.6–1.0)	−70	−45	15	20

Control Effect

Based on the above control strategy, the vehicle model is built by using the Simulink module in MATLAB software, and the control strategy is written in the Stateflow module. Vehicle model parameters set reference to the parameters of the Jetta car, front and rear wheel braking force distribution coefficient is 0.4, and simulation step is 0.01 s, the vehicle starts braking at 0.5 s after the initial speed, the high adhesion coefficient road is asphalt road, the adhesion coefficient is 0.8; the low is ice and snow road, the adhesion coefficient is 0.2. At the speed of 80 km/h, the two kinds of road conditions, control curves and analysis, reflect the situation of one front wheel. In the first 2 s, the road is the low adhesion coefficient road, after 2 s is the high.

The braking simulation curve is shown in Fig. 3. At the time of 2 s, because of the increase in the road adhesion coefficient, a sharp decline in the slip rate, resulting in the brake force, is much less to achieve the optimal slip rate at this time; after the ABS control about 1 s, the optimal slip rate is achieved, which satisfies the control requirement of ABS, Fig. 3a. At all stages of braking, except near the contact point of docking road, wheel speed is steadily increased or decreased, to meet the requirements of vehicle's stability, Fig. 3b.

In the first 2 s, the road is the high adhesion coefficient road, after 2 s is the low adhesion coefficient road. The simulation results are shown in Fig. 4; compared to Fig. 3, the results are both similar and opposite.
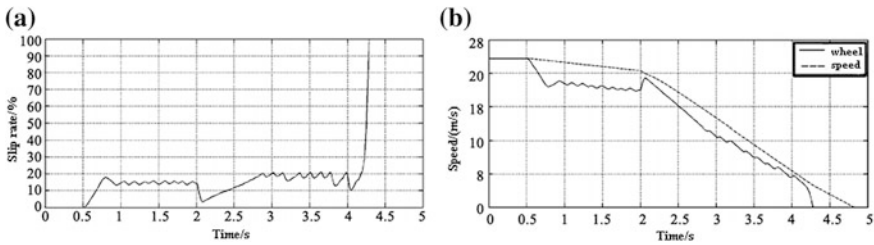
Fig. 3 ABS control simulation curves from low road adhesion coefficient to high road adhesion coefficient
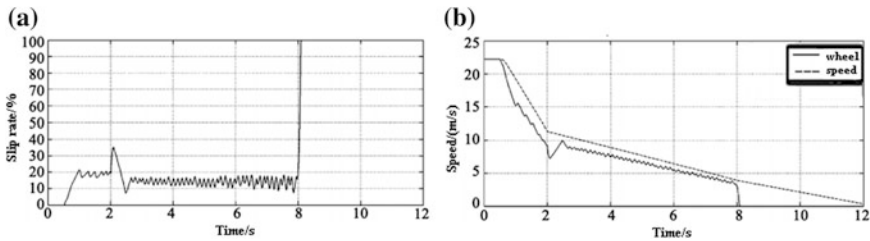
Fig. 4 ABS control simulation curves from high road adhesion coefficient to low road adhesion coefficient

From Figs. 3 and 4, different road conditions can be seen; ABS control of the self-adjustment threshold value can control the longitudinal stability of the vehicle in a stable range and ensure slip rate near the current optimal slip rate for the small-scale fluctuations; the wheel speed is stable and braking efficiency is improved, which satisfy the control requirement of ABS.

Conclusions

In this paper, we built a 3-DOF vehicle model based on the MATLAB software, the ABS logic threshold control algorithm is simulated in Simulink and Stateflow modules. Through the simulation of the ABS self-adjusting threshold algorithm under different typical operating conditions, the validity of the self-adjusting threshold algorithm is verified in this paper. The self-adjusting threshold algorithm not only can effectively control the speed and maintain the stability of wheel speed, but also can improve the control effect of ABS, which achieve the desired effect of the system can effectively control the vehicles in all conditions.

References

1. Zheng, Taixiong, and Fulei Ma. 2011. Research on vehicle ABS logic threshold control method. *Journal of Chongqing University of Posts and Telecommunications* 23 (4): 455–457.
2. Zhu, Weixing, Yinchang Chen. 2004. Application and Simulation of fuzzy PID control in vehicle ABS. *Journal of Jiangsu University* 25 (4): 311–313.
3. Layne, J.R., Passino, K. M., and S. Yurkovich. 1993. Fuzzy learning control for anti-skid braking systems. *IEEE Transactions on Control Systems Technology* 1 (2): 122–129.
4. Lin, C.M., and C.F. Hsu. 2003. Self-learning fuzzy sliding-mode control for antilock braking systems. *IEEE Transact ions on Control Systems Technology* 11 (2): 273–278.
5. Jiang, F.J., and Z.Q. Gao. 2001. An application of nonlinear PID control to a class of truck ABS problems. *IEEE Process Decision and Control* 1 (1): 516–521.
6. Cabrera, J.A., Ortiz, A.., Castillo, J. J. et al. 2005. A fuzzy logic control for antilock braking system integrated in the IMMa tire test bench. *IEEE Transactions on Vehicular Technology* 54 (6): 1937–1949.
7. Xiong, Jian. 1997. Signal analysis and threshold select of anti-lock braking control. *Journal of Yunnan University of Technology* 13 (2): 4–5.
8. Weida, Wang. 2008. Research on control method of ABS/ASR/DYC integrated system for vehicle. Beijing: Beihang University.

Failure Analysis of Metro Door System Based on Fuzzy TOPSIS

Xiao-qing Cheng, Yong Qin, Li-min Jia and Zheng-yu Zhang

Abstract Door system is a key subsystem of metro train, and conducting failure analysis to determine the failure modes which have great criticality on door system can be helpful for improving door reliability and perfecting maintenance decision. There being ambiguity in evaluating risk factors, the fuzzy theory is introduced into the TOPSIS to build the fuzzy TOPSIS method. Firstly, to conduct fuzzy evaluation on risk factors of each failure mode and to obtain weighted decision matrix, then the closeness coefficient of each failure mode could be obtained by the weighted decision matrix; finally, by conducting failure analysis on metro door system using this method, it can be concluded that the modes 'EDCU function is broken' 'Nut component wears out' and 'Limit switch S1 wears out' have great criticality on door system. The result is consistent with maintenance experts in the field and could provide technical support for design and maintenance of door system.

Keywords Urban rail transit · Failure analysis · Fuzzy TOPSIS · Metro door · Criticality

X. Cheng (✉) · Y. Qin · L. Jia · Z. Zhang
State Key Laboratory of Rail Traffic Control and Safety,
Beijing Jiaotong University, Beijing 100044, China
e-mail: xqcheng@bjtu.edu.cn

X. Cheng · Y. Qin · L. Jia · Z. Zhang
School of Traffic and Transportation, Beijing Jiaotong University,
Beijing 100044, China

X. Cheng
Beijing Key Laboratory of Performance Guarantee on Urban Rail
Transit Vehicles, Beijing 100044, China

Introduction

For metro train, door system is one of the key subsystems, and its high failure rate has attracted the door design department and metro operating company great attention. Failure analysis has become an important link of modern product development and maintenance; it is imperative to conduct failure analysis of door system on the purpose of helping metro company to make reasonable maintenance decision.

Chinese and foreign scholars have made many efforts to make failure analysis for identifying those failure modes of great criticality. Shang [1] applied FMECA based on fuzzy comprehensive evaluation to failure analysis of automatic loading transmission system. Li [2] utilized the fuzzy TOPSIS method (Technique for Order Performance by Similarity to Ideal Solution) to implement criticality analysis on failure modes of processing operations of a product. Zhu [3] employed FTA to conduct failure analysis of metro door system, and the key failure modes were found.

Up to now, any research literatures applying fuzzy TOPSIS method to conduct failure analysis on metro door system have not retrieved. So the fuzzy TOPSIS is applied to failure analysis of door system. On the basis of fuzzy assessment on risk factors, the fuzzy TOPSIS is used to obtain the closeness coefficient of each failure mode, and then the criticality sequence of failure modes is determined. The result could provide technical support for door maintenance decision.

Principle and Mathematical Model of Fuzzy TOPSIS

In traditional failure criticality analysis, an exact score value is assigned to occurrence (O), severity (S), and detectability (D). But it is very difficult to assign exact values to the three risk factors because of fuzziness and uncertainty in failure analysis. Using the fuzzy theory, appropriate linguistic terms and fuzzy numbers could be applied to replace the exact value to assess risk factors; the problem of fuzziness and uncertainty in the failure analysis will be solved.

The TOPSIS method is one of the classical multicriteria decision making methods, which was developed by Hwang and Yoon [4]. It is based on the concept that the chosen alternative should have the shortest distance from the positive ideal solution and the farthest from the negative ideal solution. It can overcome the problem of risk factors not being entitled to weights, which provide scientific reference for failure criticality analysis. So the fuzzy TOPSIS theory is built for analyzing the criticality of door failure modes. The principle diagram of fuzzy TOPSIS is shown in Fig. 1.

Fig. 1 Principle diagram of fuzzy TOPSIS

| Construct the fuzzy decision matrix | → | Construct the weight fuzzy decision matrix | → | Obtain the closeness coefficient | → | Confirm the criticality sequence |

Linguistic terms	Fuzzy score
Very low (VL)	(1, 1, 3)
Low (L)	(1, 3, 5)
Medium (M)	(3, 5, 7)
High (H)	(5, 7, 9)
Very high (VH)	(7, 9, 10)
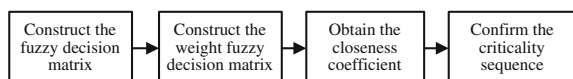
Table 1 Fuzzy evaluation scores for failure modes

Construct the Fuzzy Decision Matrix

The fuzzy TOPSIS is introduced in this paper on the purpose of making the analysts to assess risk factors conveniently, which could avoid the defect of the exact value assessment on risk factors and make failure analysis more reasonably and feasibly. Table 1 gives the linguistic scale for evaluating the ratings of failure modes with respect to three risk factors O, S, and D [5].

Assuming the failure analysis team has K analysts, $X = \{x_1, x_2, \ldots, x_n\}$ represents failure modes set and $A = \{a_1, a_2, a_3\}$ denotes risk factors set, including O, S, and D. $\tilde{r}_{ij}$ is the rating of failure mode x_i with respect to risk factor a_j, and then the fuzzy decision matrix for n failure modes is expressed as [6]:

$$\tilde{r}_{ij} = \frac{1}{K}(\tilde{r}_{ij}^1 + \tilde{r}_{ij}^2 + \cdots + \tilde{r}_{ij}^K) \tag{1}$$

$$\tilde{R} = \begin{bmatrix} \tilde{r}_{11} & \tilde{r}_{12} & \tilde{r}_{13} \\ \tilde{r}_{21} & \tilde{r}_{22} & \tilde{r}_{23} \\ \vdots & \vdots & \vdots \\ \tilde{r}_{n1} & \tilde{r}_{n2} & \tilde{r}_{n3} \end{bmatrix} \tag{2}$$

Supposing $\tilde{r}_{ij} = (a_{ij}, b_{ij}, c_{ij})$, when A is the set of cost criteria, the normalized fuzzy decision matrix is obtained by Eqs. (3) and (4) [7].

$$\tilde{b}_{ij} = \left(\frac{a_{ij}}{c_j^+}, \frac{b_{ij}}{c_j^+}, \frac{c_{ij}}{c_j^+} \right) \tag{3}$$

$$c_j^+ = \max_i c_{ij} \tag{4}$$

Construct the Fuzzy Weighted Decision Matrix

When evaluating weights of risk factors, analysts' opinions are usually subjective and ambiguous, so the fuzzy AHP method is introduced in this paper to determine

the weights of three risk factors. To make analysts express their subjective opinions comfortably, the triangular fuzzy numbers are used to represent the relative importance scale of risk factors, as shown in Table 2 [8].

The judgment matrix:

$$\tilde{A} = [\tilde{a}_{ij}]_{3\times3} = \begin{bmatrix} 1 & \tilde{a}_{12} & \tilde{a}_{13} \\ \tilde{a}_{21} & 1 & \tilde{a}_{23} \\ \tilde{a}_{31} & \tilde{a}_{32} & 1 \end{bmatrix}$$

where $\tilde{a}_{ij} = (l_{ij}, m_{ij}, u_{ij})$; it is determined by the average method according to analysts' opinions; $\tilde{a}_{ij}^k$ denotes the rating of the kth analysts on the relative importance between the risk factors i and j. $\tilde{a}_{ij} = \frac{1}{\tilde{a}_{ji}} = \left(\frac{1}{u_{ij}} \quad \frac{1}{m_{ij}} \quad \frac{1}{l_{ij}} \right)$, $\tilde{a}_{ii} = 1$.

For matrix $\tilde{A}$, the geometric mean method is used to calculate the fuzzy weights, the weights of ith risk factors.

$$\tilde{w}_i = \tilde{r}_i \times (\tilde{r}_1 + \tilde{r}_2 + \tilde{r}_3)^{-1}$$

where $\tilde{r}_i = (\tilde{a}_{i1} \times \tilde{a}_{i2} \times \tilde{a}_{i3})^{1/3}$, i = 1, 2, 3.

For weight $\tilde{w}_i$, the centroid defuzzification method is utilized. Supposing $\tilde{w}_i = (l, m, u)$ and its membership function $u(z)$, then the equation of centroid defuzzification is as follows.

$$w = \frac{\int u(z) \cdot z \mathrm{d}z}{\int u(z) \mathrm{d}z}$$

On the basis of the normalized fuzzy decision matrix $B = (b_{ij})_{nx3}$ and risk factor weights $W = (w_1, w_2, w_3)^T$, the weighted fuzzy matrix is calculated by Eq. (5):

$$U = \begin{bmatrix} u_{11} & u_{12} & u_{13} \\ u_{21} & u_{22} & u_{23} \\ \vdots & \vdots & \vdots \\ u_{n1} & u_{n2} & u_{n3} \end{bmatrix} = \begin{bmatrix} b_{11}w_1 & b_{12}w_2 & b_{13}w_3 \\ b_{21}w_2 & b_{22}w_2 & b_{23}w_3 \\ \vdots & \vdots & \vdots \\ b_{n1}w_1 & b_{n2}w_2 & b_{n3}w_3 \end{bmatrix} \tag{5}$$

	Fuzzy scale	Linguistic terms	Fuzzy number
Table 2 Fuzzy numbers for the relative importance of risk factors	$\tilde{1}$	Fairly important	(1, 1, 2)
	$\tilde{3}$	Slightly important	(2, 3, 4)
	$\tilde{5}$	Obviously important	(4, 5, 6)
	$\tilde{7}$	Highly important	(6, 7, 8)

Obtain the Closeness Coefficient

The positive ideal solution X^+ and negative ideal solution X^- in this paper are defined as [9]:

$$X^+ = \left(M_1^+, M2^+, M3^+\right), M_1^+ = M_2^+ = M_3^+ = (0, 0, 0)$$

$$X^- = \left(M_1^-, M2^-, M3^-\right), M_1^- = M_2^- = M_3^- = (1, 1, 1)$$

The distance of each failure mode from X^+ and X^- can be currently calculated by Eqs. (6) and (7).

$$d_i^+ = \sum_{j=1}^{3} d(u_{ij}, M_j^+) \tag{6}$$

$$d_i^- = \sum_{j=1}^{3} d(u_{ij}, M_j^-) \tag{7}$$

where $d(A, B)$ is the distance measurement between two fuzzy numbers, calculated using the following formula.

$$d(A, B) = \sqrt{\frac{1}{3}[(a_1 - b_1)^2 + (a_2 - b_2)^2 + (a_3 - b_3)^2]} \tag{8}$$

where $A = (a_1, a_2, a_3)$ and $B = (b_1, b_2, b_3)$ are two triangular fuzzy numbers. A closeness coefficient is defined to determine the ranking order of all failure modes; the closeness coefficient of each alternative is calculated as:

$$d_i = \frac{d_i^-}{d_i^- + d_i^+} \tag{9}$$

Confirm the Criticality Sequence

All tables must be centered in the column and numbered consecutively (in Arabic numbers). Table headings should be placed above the table. Place tables as close as possible to where they are mentioned in the main text (see Tables 1 and 2).

The higher closeness coefficient reflects the smaller criticality of the identified failure modes. We could sort the coefficient by the ascending order and confirm the criticality sequence, and formulate some targeted improvement measures.

An Illustrative Example

A statistical analysis of door system fault information which was found in the main line operation and maintenance in this paper is carried out, and ten failure modes of high failure rate are selected for failure analysis, as shown in Table 3.

Suppose the criticality analysts are comprised of five experienced analysts who come from design, assembly, and maintenance departments, they are given the same weight. The assessment information of three risk factors and weights provided by the five criticality analysts is presented in Table 4.

Table 3 Common failure modes of door system

No.	Failure modes
FM1	Maladjustment of pinch roller clearance
FM2	Trolley guiding wears out
FM3	Portable door frame have cracks
FM4	Disability of EDCU
FM5	Limit switch $S1$ wears out
FM6	Disability of limit switch $S3$
FM7	Limit Switch $S4$ wears out
FM8	Loose of screw support
FM9	Nut component wears out
FM10	Poor lubrication of middle unlocking component

Table 4 Fuzzy assessments of risk factors for door failure modes

Risk factors	Analysts	Failure modes									
		FM1	FM2	FM3	FM4	FM5	FM6	FM7	FM8	FM9	FM10
Occurrence	TM1	M	L	M	M	M	L	L	L	M	L
	TM2	L	M	M	H	M	L	L	L	M	M
	TM3	L	L	M	M	M	L	L	L	M	L
	TM4	L	L	M	M	M	L	L	L	M	L
	TM5	L	L	L	M	M	VL	L	L	H	L
Severity	TM1	H	L	L	H	H	H	H	L	H	L
	TM2	M	M	M	H	M	M	M	M	M	L
	TM3	M	L	L	VH	H	H	H	L	H	L
	TM4	H	M	L	H	H	H	H	L	H	L
	TM5	M	M	M	H	H	H	H	L	H	L
Detection	TM1	M	L	L	M	M	H	M	L	M	M
	TM2	M	M	M	H	M	M	M	M	L	M
	TM3	H	M	L	M	M	M	M	M	M	M
	TM4	M	L	M	H	M	H	M	M	M	M
	TM5	M	M	M	M	H	H	H	M	M	M

According to Eq. (1), group assessment information of ten failure modes is shown in Table 5. Moreover, fuzzy decision matrix of ten failure modes on the basis of Eq. (2) is as follows:

$$\tilde{R} = \begin{bmatrix} (1.4, 3.4, 5.4) & (3.8, 5.8, 7.8) & (3.4, 5.4, 7.4) \\ (1.4, 3.4, 5.4) & (2.2, 4.2, 6.2) & (2.2, 4.2, 6.2) \\ (2.6, 4.6, 6.6) & (1.8, 3.8, 5.8) & (2.2, 4.2, 6.2) \\ (3.4, 5.4, 7.4) & (5.4, 7.4, 9.2) & (3.8, 5.8, 7.8) \\ (3.0, 5.0, 7.0) & (4.6, 6.6, 8.6) & (3.4, 5.4, 7.4) \\ (1.0, 2.6, 4.6) & (4.6, 6.6, 8.6) & (4.2, 6.2, 8.2) \\ (1.0, 3.0, 5.0) & (4.6, 6.6, 8.6) & (3.4, 5.4, 7.4) \\ (1.0, 3.0, 5.0) & (1.4, 3.4, 5.4) & (2.6, 4.6, 6.6) \\ (3.4, 5.4, 7.4) & (4.6, 6.6, 8.6) & (2.6, 4.6, 6.6) \\ (1.4, 3.4, 5.4) & (1.0, 3.0, 5.0) & (3.0, 5.0, 7.0) \end{bmatrix}$$

The fuzzy AHP method is used to determine the weights of risk factors. In terms of judgment matrix given by five analysts, fuzzy weights of three risk factors are determined, respectively.

$$\tilde{w}_O = (0.17, 0.31, 0.64)$$

$$\tilde{w}_S = (0.26, 0.53, 0.99)$$

$$\tilde{w}_D = (0.08, 0.15, 0.32)$$

After defuzzification, weights can also be normalized: w = (0.32, 0.51, 0.17). The fuzzy decision matrix is normalized by Eqs. (3) and (4), and in terms of Eq. (5) and the weights w, the fuzzy weighted decision matrix of ten failure modes is as follows.

Table 5 Group assessment of failure modes

Failure modes	Occurrence	Severity	Detection
FM1	(1.4, 3.4, 5.4)	(3.8, 5.8, 7.8)	(3.4, 5.4, 7.4)
FM2	(1.4, 3.4, 5.4)	(2.2, 4.2, 6.2)	(2.2, 4.2, 6.2)
FM3	(2.6, 4.6, 6.6)	(1.8, 3.8, 5.8)	(2.2, 4.2, 6.2)
FM4	(3.4, 5.4, 7.4)	(5.4, 7.4, 9.2)	(3.8, 5.8, 7.8)
FM5	(3.0, 5.0, 7.0)	(4.6, 6.6, 8.6)	(3.4, 5.4, 7.4)
FM6	(1.0, 2.6, 4.6)	(4.6, 6.6, 8.6)	(4.2, 6.2, 8.2)
FM7	(1.0, 3.0, 5.0)	(4.6, 6.6, 8.6)	(3.4, 5.4, 7.4)
FM8	(1.0, 3.0, 5.0)	(1.4, 3.4, 5.4)	(2.6, 4.6, 6.6)
FM9	(3.4, 5.4, 7.4)	(4.6, 6.6, 8.6)	(2.6, 4.6, 6.6)
FM10	(1.4, 3.4, 5.4)	(1.0, 3.0, 5.0)	(3.0, 5.0, 7.0)

Table 6 Criticality sequences of door system failure modes

Failure modes	Closeness coefficient	Criticality sequence
FM1	0.800	6
FM2	0.834	8
FM3	0.824	7
FM4	0.741	1
FM5	0.762	3
FM6	0.787	4
FM7	0.788	5
FM8	0.848	9
FM9	0.760	2
FM10	0.848	9

$$\tilde{U} = \begin{bmatrix} (0.06, 0.15, 0.23) & (0.21, 0.32, 0.43) & (0.07, 0.11, 0.15) \\ (0.06, 0.15, 0.23) & (0.12, 0.23, 0.34) & (0.05, 0.09, 0.13) \\ (0.11, 0.20, 0.29) & (0.10, 0.21, 0.32) & (0.05, 0.09, 0.13) \\ (0.15, 0.23, 0.32) & (0.30, 0.42, 0.51) & (0.08, 0.12, 0.16) \\ (0.13, 0.22, 0.30) & (0.26, 0.37, 0.48) & (0.07, 0.11, 0.15) \\ (0.04, 0.11, 0.20) & (0.26, 0.37, 0.48) & (0.09, 0.13, 0.17) \\ (0.04, 0.13, 0.22) & (0.26, 0.37, 0.48) & (0.07, 0.11, 0.15) \\ (0.04, 0.13, 0.22) & (0.08, 0.19, 0.30) & (0.05, 0.10, 0.14) \\ (0.15, 0.23, 0.32) & (0.26, 0.37, 0.48) & (0.05, 0.10, 0.14) \\ (0.06, 0.15, 0.23) & (0.06, 0.17, 0.28) & (0.06, 0.10, 0.15) \end{bmatrix}$$

According to Eqs. (6)–(9), the closeness coefficients of 10 failure modes and the criticality sequences are shown in Table 6.

From Table 6, it can be apparently concluded that the closeness coefficients of 4 (Disability of EDCU), 9(Nut component wears out), and 5(Limit switch $S1$ wears out) are far less than the coefficients of other failure modes, and they should be regarded as the key failure modes and be the focus of attention in the routine maintenance. In addition, reasonable improvement measures should also be given to failure mode 6(Disability of Limit switch $S3$), 7(Limit switch $S4$ wears out), 1 (Maladjustment of pinch roller clearance), 3(Portable door frame have cracks), 2 (Trolley guiding wears out), 8(Loose screw support), 10(Poor lubrication of middle unlocking component).

Conclusion

It is more rational and applicable to apply the fuzzy TOPSIS method to make failure analysis, and the method could be used as an useful supplement to the traditional failure analysis methods. By conducting failure analysis on metro door system in

this paper using the fuzzy TOPSIS method, it has been proved that the three failure modes, 'Disability of EDCU', 'Nut component wears out' and 'Limit switch $S1$ wears out,' have great criticality on door system and should be the focus of maintenance department. The obtained result is consistent with the maintenance engineers and will provide the decision-making suggestions for the door design improvement and maintenance.

References

1. Shang, W.R., Zj Dai, and C.Z. Hou. 2006. FMECA based on fuzzy evaluation for embedded system. *Computer Simulation* 23 (10): 53–56.
2. Li, H.F., Y. Ren, and S.K. Zeng. 2008. Decision making method for failure mode harmfulness based on TOPSIS. *O.I. Automation* 27 (11): 90–92.
3. Zhu, X.J., J.B. Wang., and Z.M. Yin. 2006. On the reliability compartment door in Shanghai metro trains. *Urban Mass Transit* 3, 31–34.
4. Kutlu, A.C., and M. Ekmekcioglu. 2012. Fuzzy failure modes and effects analysis by using fuzzy TOPSIS-based fuzzy AHP. *Expert Systems with Applications* 39, 61–67.
5. Chen, C.T. 2000. Extensions of the TOPSIS for group decision-making under fuzzy environment. *Fuzzy Sets and Systems* 114, 1–9.
6. Li, J.Q., J.X. Yu, and C. Miao. 2012. Talent selection model based on fuzzy AHP and fuzzy TOPSIS. *Journal of Liaoning Shihua University* 32 (4): 88–91.
7. Wang, T.C., and T.H. Chang. 2007. Application of TOPSIS in evaluating initial training aircraft under a fuzzy environment. *Expert Systems with Applications* 33: 870–880.
8. Wang, R.J. 2010. Evaluation of knowledge absorptive capability of enterprise based on fuzzy AHP and TOPSIS. *Industrial Engineering Journal* 13 (3): 105–110.
9. Chen, J., H.S. Zhang, and J.G. Zhang. 2010. Performance evaluation of missile weapon systems based on advanced TOPSIS theory. *Computer Simulation* 27 (9): 83–87.

The Gradation Relationship Model and Application of Urban Multimodal Transit Networks

Xian-cai Jiang, Jian-liang Mo and Ya-hui Li

Abstract Aiming at the situation that the rational allocation of multimodal transit networks planning is ambiguous in China, firstly, the characteristics of residents' transit trip distance are analyzed, and the distance distribution model of residents' transit trip is built with double parameters Weibull distribution; secondly, preponderant trip distance of each transit network is presented based on the generalized travel cost; finally, the gradation model of urban multimodal transit networks is proposed based on the balanced analysis of supply and demand. Moreover, taking transit networks of Harbin as an example, a reasonable scale and gradation relationship of each transit network of Harbin in 2020 is calculated; the results are matched with the expected demand in 2020, which demonstrates that the model is reasonable and feasible. This paper provides a new approach for determining the reasonable gradation relationship of multimodal transit networks.

Keywords Multimodal transit networks · Dominance matrix · Trip distance distribution · Grade configuration

Introduction

With the rapid development in China's economy, urban traffic demand increases sharply, which has led to traffic congestion, traffic accidents, and other issues. Throughout the history of urban traffic management, the experience of some developed countries illustrates that it is the fundamental way to solve the increasingly serious problem of urban traffic through developing urban public transport and establishing safe, reliable, and energy-efficient urban public transport system.

X. Jiang (✉) · J. Mo · Y. Li
School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, China
e-mail: jxc023@126.com

The multimodal public transport system refers to the generic name of the urban public transport system which is composed of two or more forms of public transport. The construction of China's urban public transport system is in a period of rapid development; it is essential for laying the dominance of urban public transport system to establish the function explicit, the grade clear multimodal transit networks.

In recent years, the researches mostly focus on the optimization of transit network and the connection of multimodal public transport, but less attention has been paid to the grade configuration of transit network. Ciaffi et al. [1] took the maximum of network service range and the minimum of total network travel time as the objective function and proposed the genetic algorithm to get the best route. Saeed et al. [2] who took the investment budget, the restrictions of the service level, and other factors as the constraint conditions put forward the heuristic algorithm to optimize the route candidate set which was created by gravity model and shortest path model. Stefaan et al. [3] introduced trip chains to analyze the relationship the choice of transport mode and time factors, and provided the theoretical basis for the analysis of multimodal public transport demand. Ferber et al. [4] analyzed the network survivability of 14 cities' public transit network all over the world under random attacks and deliberate attacks and revealed rich behavior beyond that of the ubiquitous scale-free complex network. Chen et al. [5] established the model of travel time reliability, which provided the support to improve the service level of multimodal transit networks. Zhang et al. [6] analyzed the topological characteristics and functional properties of Shanghai metro network in order to access the robustness and reliability according to the complex network theory. When multimodal transit network planning is carried out currently, the determination of the construction scale of transit networks of the various modes in China usually adopt analogy method [7, 8]. That is, drawing lessons from already built multimodal transit networks scale at home and abroad; then, the transit network length of various modes is estimated according to the form, area, GDP, population, etc. of these cities, and the right range of construction scale in the target city is determined finally. However, the method ignores the differences of traffic policy and traffic structure among different cities, so the calculated scale may not be reasonable.

Throughout the development process of each urban road network, it has also gone from haphazard to function explicit, grade clear. The preponderant interval of each grade road and the sharing ratio of reasonable traffic volume are determined through the analysis of characteristics, including travel distance, mode, and speed based on the idea of the balance of supply and demand, starting from the road function [9]. Transit network and road network belong to network system, and there is a certain similarity. Therefore, the paper draws on the grade configuration idea of the balance of supply and demand of urban road network, and determines the reasonable scale of each transit network.

The Model of Residents' Transit Trip Distance Distribution

Residents' transit trip distance distribution data of Shenzhen [10] and Harbin City are shown in Tables 1 and 2. In order to more vividly describe the characteristics of the data presented, they are plotted in Figs. 1 and 2.

In Figures 1 and 2, it can be seen that the residents' transit trip distance distribution shows similar Gamma function characteristics and can be fitted with the double parameters Weibull distribution, and its standard distribution function is shown in Eq. (1):

$$F(x; \alpha, \beta) = 1 - \exp\left(-\left(\frac{x}{\beta}\right)^{\alpha}\right) \tag{1}$$

wherein α is the shape parameter and β is the proportion parameter. Under the same shape parameter condition, the larger the β is, the flatter the curve will be. Under the same proportion parameter condition, the larger the α is, the sharper the curve will be.

Taking advantage of the model to fit the data of residents' transit trip distance distribution in Shenzhen and Harbin, we can obtain the corresponding fitting curves as shown in Figs. 3 and 4.

The fitting results are shown in Table 3.

As can be seen from Table 3, the effect which adopts double parameters Weibull distribution to fit the data of urban residents' transit trip distance is significant.

Table 1 Data of residents' transit trip distance distribution in Shenzhen

Travel distance (km)	Proportion (%)	Travel distance (km)	Proportion (%)
≤ 1	2.17	≤ 15	88.44
≤ 3	23.84	≤ 20	91.61
≤ 6	57.34	≤ 30	94.61
≤ 10	80.84	≤ 50	96.11

Table 2 Data of residents' transit trip distance distribution in Harbin

Travel distance (km)	Proportion (%)	Travel distance (km)	Proportion (%)	Travel distance (km)	Proportion (%)
≤ 1	5.39	≤ 7	77.48	≤ 13	98.21
≤ 2	17.92	≤ 8	84.58	≤ 14	98.76
≤ 3	32.82	≤ 9	89.31	≤ 15	99.19
≤ 4	45.87	≤ 10	92.38	≤ 16	99.55
≤ 5	57.61	≤ 11	94.71	≤ 17	99.70
≤ 6	68.32	≤ 12	96.78	≤ 18	99.84
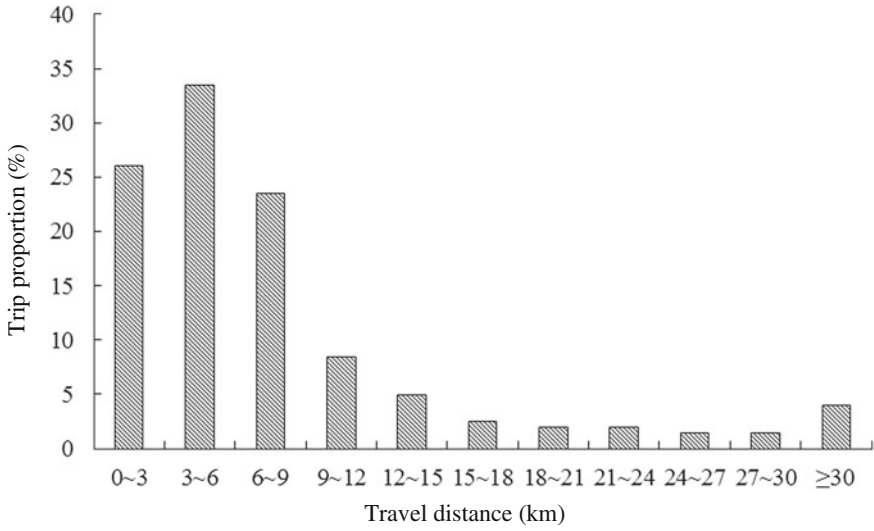
Fig. 1 Residents' transit trip distance distribution in Shenzhen
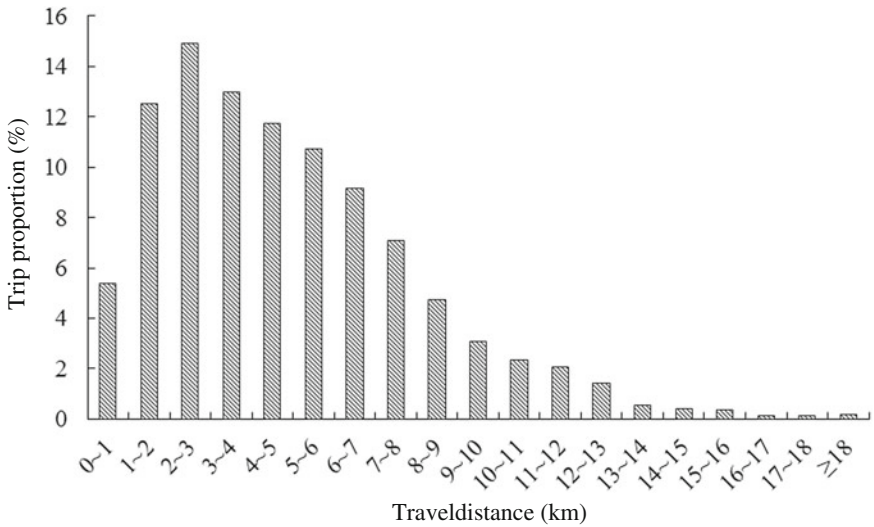
Fig. 2 Residents' transit trip distance distribution in Harbin

Fig. 3 Fitting chart of double parameters Weibull distribution in Shenzhen
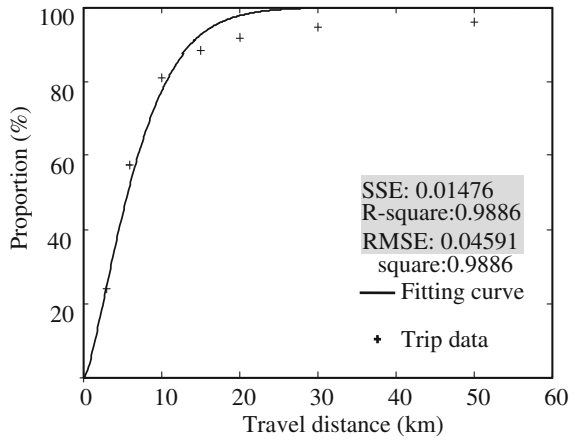
Fig. 4 Fitting chart of double parameters Weibull distribution in Harbin
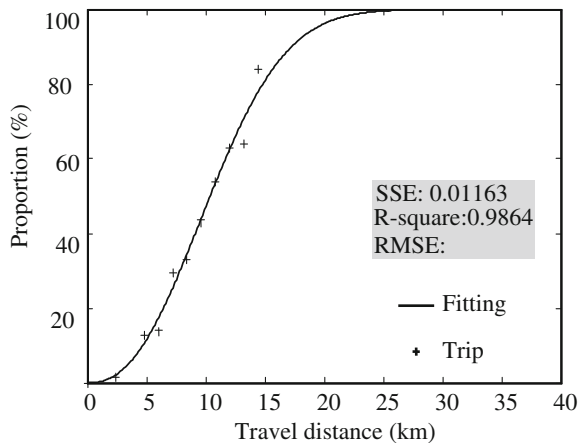
Table 3 Fitting model parameter and correlation coefficient

City	Model parameter	Correlation coefficient
Shenzhen	$\alpha = 1.327$ $\beta = 7.315$	0.9886
Harbin	$\alpha = 1.592$ $\beta = 5.463$	0.9998

The Model of Preponderant Trip Distance of Each Transit Network

The Generalized Cost of Residents' Transit Trip

In the field of transportation, the generalized cost refers to the comprehensive evaluation of all the attributes of transportation service level in the transportation network planning, and it is also an important part of reflecting the supply grade in the transportation economy [11].

The generalized cost of residents' transit trip is the sum of all cost to complete once trip by bus, and generally is composed of measured directly economic cost and measured indirectly economic cost. The measured directly economic cost generally refers to take a bus to pay the ticket fare. The measured indirectly economic cost is the focus of reflecting generalized cost function, including bus travel time, comfort, and security.

1. Bus travel time

The time of resident once trip by bus is mainly composed of three parts: walking time, waiting time, and travel time; the bus travel time is defined as Eq. (2):

$$T = T_{\text{walk}} + T_{\text{wait}} + T_{\text{travel}} \tag{2}$$

where T_{walk} is walking time (h), T_{wait} is waiting time (h), and T_{travel} is travel time (h).

2. Ticket fare

The ticket fare is the factor of identifying and quantifying easily in various travel factors. Under no significant differences of other influential factors, the condition, and same distance, people are more inclined to choose low-fare trip mode. Residents' transit travel fare is formulated as Eq. (3) (considering only bus transit, BRT, and metro):

$$C = \theta_1 \cdot (n_1 + 1) \cdot \lambda_1 + \theta_2 \cdot d_2 \cdot \lambda_2 + \theta_3 \cdot d_3 \cdot \lambda_3 \tag{3}$$

where C is the cost of resident once trip by bus (yuan); $\theta_1, \theta_2, \theta_3$ are the choice factors of bus transit, BRT, and metro, respectively, $\theta_i \in (0, 1)$, if resident takes a bus transit or metro in once trip, then $\theta_1 = \theta_3 = 1$, $\theta_2 = 0$, n_1 is the transfer times in once trip; λ_1 is the ticket fare by bus transit (yuan/person), λ_2, λ_3 are the rate of BRT and metro(yuan/person km), respectively, d_1, d_2 are travel distance of BRT and metro in once trip (km), respectively.

3. Safety

The generalized cost of safety is generally expressed through casualties index of turnover per unit passenger [12]. Despite bus transit spontaneous event reported by the media in recent years, but it is only one example. Compared with other traffic modes, the safety of metro, BRT, and bus transit is very high in this paper; therefore, the difference of residents' transit travel choice behavior caused by the safety is very small. So the safety factor is not taken into consideration.

4. Comfort

The impact of comfort on the choice of people travel mode is obvious, under the allowed time and the accepted cost conditions, and people prefer to choose traffic mode of comfortable traveling environment. With the improvement of people's living standard, the requirement for comfort is also higher and higher. However, the

quantification of comfort is a difficult problem and generally takes the ticket fares 5–10% as the monetary performance of the comfort in the application [13].

In a word, the generalized travel cost of resident bus is shown as Eq. (4):

$$F = \eta T + (1 + a)C \tag{4}$$

where F is the generalized travel cost of resident bus (yuan), η is passengers time value (yuan/h), and a is the comfort coefficient.

Wherein, the travel time value is the fee which traveler is willing to pay more to save the per unit travel time. At present, the determination of the time value usually adopts income method at home and abroad [14], and it is formulated as Eq. (5):

$$\eta = \frac{In}{D \cdot H} \tag{5}$$

where In is the annual average wage of workers in the city (yuan), D is the average work days of workers in a year(day), and H is the average work length of workers in a day(h).

The Establishment of Preponderant Trip Distance Model of Each Transit Network

① When the travel distance is short, the generalized travel cost by bus transit is lower than that of the use of BRT and metro, and more advantageous. Mainly because compared with other transit network, bus transit network is dense, the required walking distance of arriving conventional bus stops is the shortest, and the ticket fare is the lowest, meanwhile the delays caused by road congestion within short distance and other factors are acceptable. Therefore, preponderant trip distance of bus transit is less than S_1. ② With the increase in travel distance, the generalized travel cost of various modes buses also increases, but the growth rate may be large or small. When it reaches S_1, the generalized travel cost of bus transit and BRT is equal; however, the generalized cost of metro is even higher than both. After this period until reaching S_2, the generalized cost of BRT is less than others, namely preponderant trip distance of BRT is between S_1 and S_2. ③ When the travel distance continues to increase and reaches S_2, the generalized travel cost of BRT and metro is equal, and thereafter the generalized travel cost of bus transit and BRT will be higher than that of metro, namely preponderant trip distance of metro is greater than S_2. The critical value of S_1 and S_2 can be obtained simultaneously through the analysis of generalized travel cost.

1. The establishment of generalized travel cost model

   ① Under the maximum use bus transit condition, the generalized travel cost is formulated as Eq. (6):

$$F_1 = \eta T_1 + (1+a)C_1 \tag{6}$$

$$T_1 = \frac{d_0}{v_0} + \bar{t}_{1\_\text{wait}} + \frac{S - d_0}{v_1} \tag{7}$$

$$C_1 = (n_1 + 1)\lambda_1 \tag{8}$$

② Under the maximum use BRT condition, the generalized travel cost is formulated as Eq. (9):

$$F_2 = \eta T_2 + (1+a)C_2 \tag{9}$$

$$T_2 = \frac{d_0}{v_0} + \bar{t}_{2\_\text{wait}} + \frac{d_1}{v_1} + \frac{S - d_0 - d_1}{v_2} \tag{10}$$

$$C_2 = (n_1 + 1)\lambda_1 + (S - d_0 - d_1)\lambda_2 \tag{11}$$

③ Under the maximum use metro condition, the generalized travel cost is formulated as Eq. (12):

$$F_3 = \eta T_3 + (1+a)C_3 \tag{12}$$

$$T_3 = \frac{d_0}{v_0} + \bar{t}_{3\_\text{wait}} + \frac{d_1}{v_1} + \frac{d_2}{v_2} + \frac{S - d_0 - d_1 - d_2}{v_3} \tag{13}$$

$$C_3 = (n_1 + 1)\lambda_1 + d_2\lambda_2 + (S - d_0 - d_1 - d_2)\lambda_3 \tag{14}$$

2. The determination of preponderant trip distance critical value.
   When $F_1 = F_2$, it can be formulated as Eq. (15):

$$S_1 = (d_0 + d_1) + \frac{(\bar{t}_{2\_\text{wait}} - \bar{t}_{1\_\text{wait}})}{\frac{v_2 - v_1}{v_1 v_2} - \frac{(1+a)\lambda_2}{\eta}} \tag{15}$$

When $F_2 = F_3$, it can be formulated as Eq. (16):

$$S_2 = (d_0 + d_1 + d_2) + \frac{(\bar{t}_{3\_\text{wait}} - \bar{t}_{2\_\text{wait}})}{\frac{v_3 - v_2}{v_2 v_3} - \frac{(1+a)(\lambda_3 - \lambda_2)}{\eta}} \tag{16}$$

The evolution trend of the generalized travel cost of the residents in different modes of public transport is shown in Fig. 5. F_{0_bus}, F_{0_BRT}, and F_{0_metro} represent the sum of the ticket fare and the average waiting time cost by taking bus, BRT, and metro, respectively. l_1, l_2, and l_3 represent the growth of the generalized cost using only walk and bus, walk and BRT, walk and metro to travel, respectively. l_4 represents the trip mode combination with the travel distance, the lowest
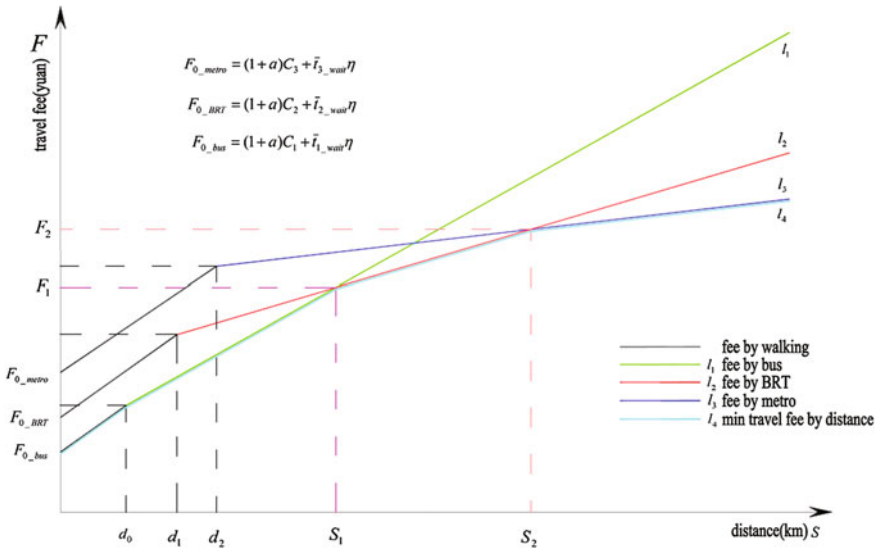
Fig. 5 Evolution trend of the generalized travel cost of the residents in different modes of public transport

generalized travel cost. d_0, d_1, and d_2 represent walking distance from the starting point to the regular bus station, BRT station, and metro station, respectively.

The Model of Urban Multimodal Transit Networks Grade Configuration

The grade configuration of transit network is divided based on the technical and economic characteristics of different bus forms and is effective means to determine a reasonable scale of each grade transit network and achieve a reasonable distribution of public transport resources.

Assuming that the attraction of some mode bus within the preponderant trip distance to passengers is significantly higher than that of other bus modes, the proportion of people choosing the mode bus to travel is higher than that of other bus modes. The paper proposes dominance matrix to characterize the advantage of different modes public transport within the travel distance, and it is shown in Table 4 and Fig. 6.

where w_{ij} is the ratio of the amount of bus travel of i mode within j distance segment accounting for the total amount, a_1, a_2, and a_3 are the trip proportions within the short distance, middle distance, and long distance, respectively b_1, b_2, and b_3 are the trip proportions of bus transit, BRT, and metro, respectively.

Table 4 Dominance matrix of multimodal transit networks

	Short distance	Middle distance	Long distance	B
Bus transit	w_{11}	w_{12}	w_{13}	$b_1 = \sum\limits_{j}^{3} w_{1j}$
BRT	w_{21}	w_{22}	w_{23}	$b_2 = \sum\limits_{j}^{3} w_{2j}$
Metro	w_{31}	w_{32}	w_{33}	$b_3 = \sum\limits_{j}^{3} w_{3j}$
A	$a_1 = \sum\limits_{i}^{3} w_{i1}$	$a_2 = \sum\limits_{i}^{3} w_{i2}$	$a_3 = \sum\limits_{i}^{3} w_{i3}$	
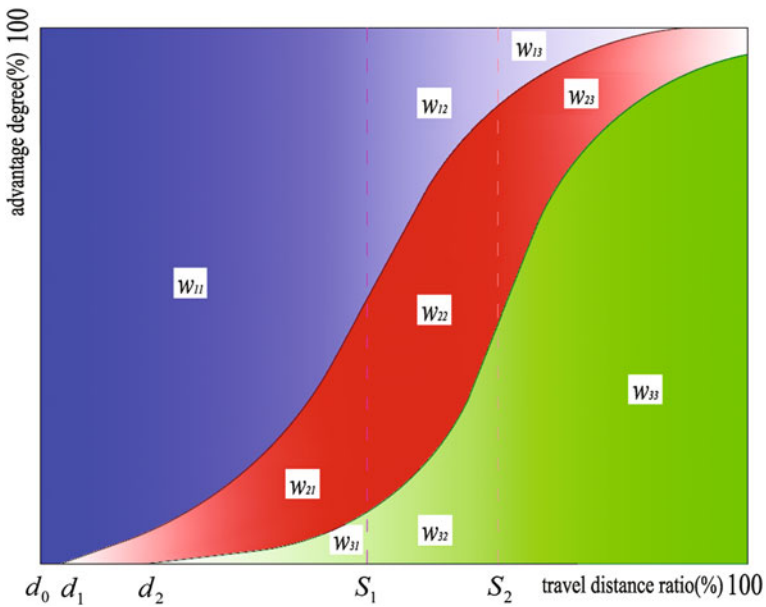
Fig. 6 Advantage of different modes of bus travel distance

1. The trip proportion within preponderant trip distance

Based on above-mentioned model of residents' transit trip distance distribution and preponderant trip distance, the trip proportion within preponderant trip distance is determined.

① The proportion of passenger travel within $[0, S_1]$ is shown as Eq. (17):

$$a_1 = \int_0^{s_1} f(s)ds = F(S_1) \tag{17}$$

② The proportion of passenger travel within $[S_1, S_2]$ is shown as Eq. (18):

$$a_1 = \int_{s_1}^{s_2} f(s)ds = F(S_2) - F(S_1) \tag{18}$$

③ The proportion of passenger travel within $[S_2, +\infty)$ is shown as Eq. (19):

$$a_3 = \int_{S_3}^{+\infty} f(s)ds = 1 - \int_0^{S_3} f(s)ds = 1 - F(S_3) \tag{19}$$

2. The trip proportion of each transit

On the premise of meeting the residents' transit demand, each transit trip proportion model which takes the minimum construction cost of bus as the objective function is established based on the dominance matrix, and it can be formulated as Eq. (20):

$$minZ = C_1 \cdot L_1 + C_2 \cdot L_2 + C_3 \cdot L_3 \tag{20}$$

where C_1, C_2, C_3 are the average investment per unit kilometer of bus transit, BRT, and metro network(yuan), respectively. L_1, L_2, L_3 are line-network length of bus transit, BRT, and metro (km), respectively.

The scale of each transit network can be presented through the total amount of bus travel, the trip proportion of each transit, and line-network load intensity. We set $c_i = (C_i \times Q \times \alpha)/\gamma_i$, and the objective function which takes the trip proportion of each bus as variables is shown as Eq. (21):

$$minz = c_1 b_1 + c_2 b_2 + c_3 b_3 \tag{21}$$

$$\text{s.t.} \begin{cases} w_{11} + w_{21} + w_{31} = a_1 \\ w_{12} + w_{22} + w_{32} = a_2 \\ w_{13} + w_{23} + w_{33} = a_3 \\ w_{11} > w_{21} > w_{31} \geq \varepsilon \\ w_{22} > w_{21} > w_{23} \geq \varepsilon \\ w_{33} > w_{32} > w_{31} \geq \varepsilon \\ \qquad w_{ij} > 0 \end{cases}$$

With MATLAB software and assuming reasonable boundary conditions, the approximate optimal solution $B = (b_1, b_2, b_3)^{\mathrm{T}}$ is obtained through the use of linear programming. Namely on the premise of meeting residents' transit trip demand, the construction cost of the transit trip proportion of each mode is minimum.

3. The model of multimodal transit networks grade configuration

The gradation of each transit network is a reasonable scale proportion of each transit network, and the previous work has determined reasonable scale of each transit network; therefore, grade configuration model is formulated as Eq. (22):

$$L_1 : L_2 : L_3 = \frac{Q \times \alpha \times b_1}{\gamma_1} : \frac{Q \times \alpha \times b_2}{\gamma_2} : \frac{Q \times \alpha \times b_3}{\gamma_3} = \frac{b_1}{\gamma_1} : \frac{b_2}{\gamma_2} : \frac{b_3}{\gamma_3} \qquad (22)$$

From the formula (22), we can know that the gradation of transit network is related to the load intensity of transit network and the trip proportion of each mode bus. Based on residents' transit trip distance distribution and preponderant trip distance, the trip proportion of each transit network which can reflect bus travel demand is obtained. The load intensity of transit network to a certain extent reflects the supply status, and its value reflects the people's subjective initiative.

Case Study

According to the comprehensive traffic planning in Harbin and taking advantage of related prediction and survey data, the reasonable scale and grade configuration of each transit network in Harbin in 2020 are determined.

The Model of Preponderant Trip Distance of Each Transit Network

According to the results of Table 3, the function of residents' transit trip distance and probability density in Harbin are formulated as Eqs. (23) and (24):

$$F(S) = 1 - e^{\left(\frac{S}{5.463}\right)^{1.592}}$$ (23)

$$f(S) = 0.106648 \times S^{0.592} \times e^{\left(\frac{S}{5.463}\right)^{1.592}}$$ (24)

The Preponderant Trip Distance of Each Transit Network in Harbin

1. Parameter value

First of all, it needs to determine the initial value of d_1 in order to calculate the preponderant trip distance of each transit network. The corresponding travel distance of the peak of the probability density curve of residents' transit trip distance distribution is the most attractive of the mode bus, and thereafter the travel proportion decreases with the increase in travel distance, meanwhile illustrates that its appeal is declining and travel advantage is more and more weak. According to the curve of the probability density, we can obtain $d_1 = 3$ km. The average walking distance of resident walking to the bus station is $d_0 = 250$ m. Based on relevant planning documents of Harbin, we can know that the average expected speed of bus transit, BRT, and metro are 15, 20, and 40 km/h in Harbin in 2020, the average train intervals in peak hour are 6, 3, and 3 min, and the average waiting time in once trip are 15, 5, and 3 min, respectively.

The average wage of on-the-job worker is 47,209 yuan in Harbin in 2013, except for the weekend 104 days and holidays 11 days, the average on-the-job days are 250 days, working length is eight hours a day, and the time value is 23.6 yuan/h.

The ticket fares of regular bus, BRT, and metro are, respectively, 1 yuan, within 4 km 1 yuan, after which 0.15 yuan/km and within 8 km 2 yuan, then 0.25 yuan/ km.

2. The results of preponderant trip distance

The critical values of preponderant trip distance of each mode bus are $S_1 = 6.4$ km and $S_2 = 10.8$ km, respectively. Therefore, we can obtain that the preponderant trip distance of bus transit is lesser than 6.4 km, the preponderant trip distance of BRT is 6.4–10.8 km, and the preponderant trip distance of metro is more than 10.8 km.

The Determination of the Travel Proportion of Each Transit Network

1. The total amount of bus travel in Harbin in 2020

According to Harbin city comprehensive transportation development "12th Five-Year Plan," the total amount of bus travel reaches 5.5 million passengers/day in Harbin in 2020.

2. The travel proportion of short distance, middle distance, and long distance

According to Eqs. (17)–(19), the proportion of bus travel within short distance, middle distance, and long distance in Harbin are $a_1 = 0.7238$, $a_2 = 0.2025$, $a_3 = 0.0737$, respectively.

At present, major cities in the domestic bus transit network load intensity is generally between 0.0003 and 0.0007 million passengers/(d km); BRT load intensity of Beijing, Curitiba, and other cities is in the 0.006–0.007 million passengers/(d km); domestic requirements early metro operators is not less than 0.007 million passengers/(d km), the country which has opened metro its load intensity between 0.006 and 0.02 million passengers/(d km). Integrated urban bus operations at home and abroad in different modes, considering Harbin Public Transport Strategy, buses, BRT, and rail transportation network, the average load intensity values are, respectively, 0.0006,0.006, and 0.01 million passengers/(d km).

3. The objective function

Through comprehensive considering public transport development strategy of Harbin, the average line-network load intensity of bus transit, BRT and metro are 0.6, 2, and 20 thousand passengers/(d km), respectively.

It has been put into operation of the metro Line 1 project in Harbin, the total length is 17.47 km, the total investment is 5.89 billion yuan, and the average cost is 0.332 billion yuan per unit kilometer. At present, it is actively planning three BRT lines in Harbin, and the investment situation is unknown. If it takes average cost of domestic BRT as standard, the average cost is 30–40 million per unit kilometer in Guangzhou, Xiamen, Shenyang, and other cities. Because the current urban road conditions are poor in Harbin, and the requirements of BRT for the road are higher, the cost of road reconstruction is relatively high. In a word, the average construction cost per unit kilometer of bus transit, BRT, and metro are 8, 40, and 350 million yuan, respectively, and the objective function and the constraint of Harbin public transit are shown as Eq. (25):

$$\min z = 733b_1 + 1100b_2 + 962b_3 \tag{25}$$

$$\text{s.t.} \begin{cases} w_{11} + w_{21} + w_{31} = 0.7238 \\ w_{12} + w_{22} + w_{32} = 0.2025 \\ w_{13} + w_{23} + w_{33} = 0.0737 \\ w_{11} \geq 0.2413 \\ w_{22} \geq 0.0675 \\ w_{33} \geq 0.0246 \\ w_{21} - w_{31} \geq 0.001 \\ w_{12} - w_{32} \geq 0.001 \\ w_{23} - w_{13} \geq 0.001 \end{cases}$$

Table 5 Dominance matrix of multimodal transit networks in Harbin

	S≤6.4	6.4<S≤10.8	S>10.8	B
Bus transit	0.4623	0.0735	0.008	0.5438
BRT	0.1742	0.1279	0.0312	0.3333
Metro	0.0873	0.0011	0.0345	0.1229
A	0.7238	0.20255	0.0737	

4. Model solution

Using the optimization tool linprog of MATLAB, we can obtain the transit trip dominance matrix of Harbin, which is shown in Table 5.

The calculation result is B = (0.5438, 0.3333, 0.1299)T. From Table 5, we can see that the proportion of bus travel is 63.87% in the short distance travel; in the middle distance travel, BRT travel ratio is 63.16%; in the long distance bus travel, metro travel ratio is 46.81%.

The construction and operation of public transportation systems in Beijing has advanced in the world; the scale of its travel and trip distance distribution has entered a stable period, so take advantage of Beijing in 2013, 2014 public transport data to check the study parameters. Statistical data show that [15] the average travel distance of bus transit in Beijing is 9.8 km, the bus transit travel proportion is 85.28% within 6 km, and bus transit travel within 6 km accounting for the total bus transit travel proportion is 56.2%; the average trip distance of metro is 17.8 km, the metro travel ratio is 58.78% above 12 km, and metro travel above 12 km accounting for the total metro travel proportion is 60.54%. From travel ratio data of each distance section, we can know that the bus transit occupies absolute dominance within 6 km, but more than 12 km the metro occupies absolute dominance. Due to the lack of distribution data of travel distance of BRT in Beijing city, and the scale is too small (only 53 km line-network length), BRT cannot be analyzed. However, from trip ratio data of bus transit and metro, we can determine that the calculation of the preponderant trip distance and preponderant trip matrix of bus transit, BRT, and metro is basically reasonable.

Discussion and Conclusion

So the reasonable line-network scale of bus transit, BRT, and metro are 4984, 305, and 67 km, respectively, and the reasonable gradation is 147:9:2.

The line-network length of bus transit is 4514 km in Harbin currently, and line-network load intensity is 760 person/(d km). The total length of metro 1 is 17.47 km, because the network has not yet been formed, its function has not been fully developed. Compared with other cities at home and abroad, the line-network load intensity of bus transit is too high, thus reflecting the insufficiency of the bus capacity input and unreasonableness of bus transit positioning. Therefore, the construction of the public transport system should be strengthened especially the BRT and metro in Harbin. At present, metro Line 2 and Line 3 of Harbin are preparing for construction; by 2020, it will form the cross and circle pattern, and the expected travel general the mileage is 65.17 km; In addition, the construction of three BRT lines relies on trial project. Because the construction of BRT lags behind in Harbin, it is not possible to compare its rationality; but from the viewpoint of the scale of bus transit and metro, the results match with expected demand in 2020 and illustrate that the gradation model of multimodal transit networks is reasonable and feasible.

This paper establishes the model of urban residents' transit trip distance distribution based on double parameters Weibull distribution and the data of urban residents' transit trip in Shenzhen and Harbin, proposes preponderant trip distance of transit network based on generalized travel cost, determines the model of multimodal transit networks gradation through taking advantage of dominance matrix and taking the minimum construction cost as an objective, and finally verifies the rationality of the model from this example of transit network in Harbin.

Future research can be studied in the area of the specific values of transit network load intensity which may be suitable for the urban characteristics. In addition, based on some ideal conditions, i.e., when the model takes the minimum construction cost as the objective function, policy is not taken into consideration.

References

1. Ciaffi, F., E. Cipriani, and M. Petrelli. 2012. Feeder bus network design problem: A new metaheuristic procedure and real size applications. *Procedia-Social and Behavioral Sciences* 54: 798–807.
2. Bagloee, S.A., and A.A. Ceder. 2011. Transit-network design methodology for actual-size road networks. *Transportation Research Part B: Methodological* 45 (10): 787–1804.
3. Walle, S.V., and T. Steenberghen. 2006. Space and time related determinants of public transport use in trip chains. *Transportation Research Part A: Policy and Practice* 40 (2): 151–162.
4. Von Ferber, C., T. Holovatch, Y. Holovatch, et al. 2009. Public transport networks: Empirical analysis and modeling. *The European Physical Journal B* 68 (2): 261–275.
5. Chen, C., Y. Xu, and X. Fu. 2010. Travel time reliability appraisal for urban multi-modal public transport network. *Journal of Transport Information and Safety* 4 (28): 27–30.

6. Zhang, J., X. Xu, L. Hong, et al. 2011. Networked analysis of the Shanghai metro network, in China. *Physica A: Statistical Mechanics and its Applications* 390 (23): 4562–4570.
7. Yang, J., and D. Zhang. 2008. Discussion on hierarchy planning method of urban public traffic network. *Railway Transport and Economy* 5 (30): 50–53.
8. Wang, Z., X. Luo. 2015. Scale calculation urban rail transit network under the condition of comprehensive transportation system. *Journal of Chang'an University* (Natural Science Edition, 2015) (35): 193–197.
9. Shi, F., and Y. Ju. 2013. Analysis on influencing factors of public transportation share: An empirical study of center Nanjing. *City Planning Review* 23: 76–84.
10. Shenzhen Municipal Engineering Consulting Center. 2003. *Research on the near future planning of conventional public transit network in Shenzhen [EB]*. http://www.docin.com/p-1025106653.html.
11. Jia, S. 2011. *Transportation economics*. Beijing: People's Communication Press.
12. Tang, J., C. Wang, and D. Zuo. 2013. Optimization model and arithmetic of region transportation structure based on optimization adjusting path. *Mathematics in Practice and Theory* 43 (13): 85–91.
13. Shi, F., L. Deng, and L. Huo. 2007. Boarding choice behavior and its utility of railway passengers. *China Railway Science* 6: 117–121.
14. Qi, T., D. Liu, and Y. Liu. 2008. A study of the traveling time cost of Beijing residents. *Journal of Highway and Transportation Research and Development* 25 (6): 144–146.
15. Beijing Transportation Research Center. 2015. *Beijing Transport Annual Report in 2015*. http://www.docin.com/p-1484581121.html?qq-pf-to=pcqq.c2c.

Study on Emergency Response Process of Metro Emergency Based on Stochastic Petri Nets

Junhui Yuan, Jie Xu, Yong Qin and Limin Jia

Abstract Stochastic Petri Nets (SPN) approach is used to model emergency response process of metro emergency in order to analyze the performance of emergency response process of metro emergency. With the method of "top down," we establish the basic models of each stage of metro emergency, then according to the actual rescue process with each stage mode of public places and transitions to construct a SPN model of the emergency response process. According to the possible state set of SPN mode, an isomorphic Markov Chain is developed for measuring and evaluating the performance of emergency response process of metro emergency, and the efficacious of the model is justified through the case of Beijing metro emergency plan response process.

Keywords Stochastic petri nets · Metro emergency · Emergency response process · Markov chain

Introduction

The subway emergency response process is an important part in the subway emergency safety management due to the tightness of the subway space which brings many uncertain factors to the subway emergency and causes inconvenience to forecast the link busy degree and execution efficiency of emergency response. There are many scholars that have carried out a lot of research on the emergency response of metro emergency [1–7], and they mainly focus on subway emergency response mechanism and system structure but they ignore the quantitative analysis

J. Yuan · J. Xu (✉) · Y. Qin · L. Jia
School of Traffic and Transportation, Beijing Jiaotong University,
Beijing 100044, China
e-mail: jxu1@bjtu.edu.cn

J. Yuan · J. Xu · Y. Qin · L. Jia
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University,
No.3 Shangyuancun, Haidian District, Beijing 100044, China

of emergency response process. The Petri nets modeling method is commendable to reflect the flow of personnel, materials, and other aspects of emergency response process of metro emergency. Fateh et al. [8] establish emergency evacuation process model based on Petri nets. Landry et al. [9] use Petri nets to establish the working flow model of the decision maker in the emergency management, and the simulation is carried out based on fuzzy rules. Sharmin et al. [10] study on the flood emergency management process based on Petri nets, and the feasibility of the model is verified by simulation. Zhong et al. [11] establish model of city emergency response system based on Petri nets, and illustrate the effectiveness of the mode by a simple example. However, in most of these studies analysis, the rationality and validity of emergency process are based on the basic Petri nets, which are not reflected by the time limit of the emergency response process.

In this paper, the SPN is selected to analyze emergency response process, and the isomorphic Markov chain is developed to calculate the stable probability of reachable markings. The time performance and operation efficiency of each link and the operation status of each stage in emergency response process are analyzed to provide the reference for the emergency management decision.

Analysis on Emergency Response Process of Metro Emergency

Subway emergency response process can be divided into four stages that early warning stage, decision-making stage, implementation stage, emergency recovery stage. There is certain relation between stages which constitute the emergency response system of metro.

Early Warning Stage Process

Early warning stage refers to the accident occurred after the subway, from the scene of the relevant personnel to the rescue department rushed to the scene of the accident process which consists of pre-handling, accident alarm, determine the level of information, and information submission, the detailed process is shown in Fig. 1. The relevant disposal of the early warning stage has provided information basis for decision-making stage.

Decision-Making Stage Process

The stage of emergency decision-making mainly includes alarm analysis, expert consultation, emergency plan analysis, and on-site information analysis; the detailed process is shown in Fig. 2. Emergency command group sent related
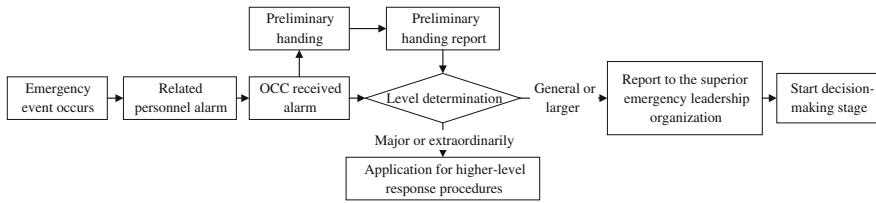
Fig. 1 Process of early warning stage
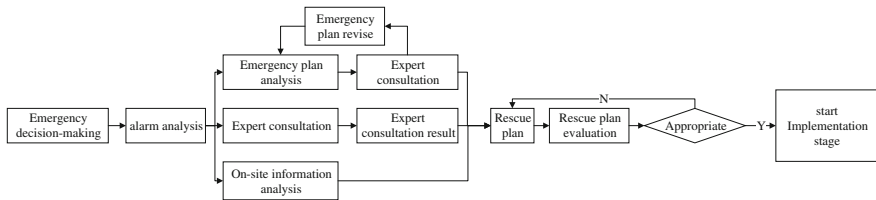
Fig. 2 Process of decision-making stage

personnel to the accident scene to investigate and analyze emergency development trend, organized experts to analyze type of the accident, influence range and emergency measures, analyzed emergency plan. Decision-making stage provides a guide for implementation stage.

Implementation Stage Process

Implementation stage is under the command of the emergency leading group, carry out a series of rescue operations according to the emergency response plan. It mainly includes on-site disposal, resource scheduling, external coordination, and information classification; the detailed process is shown in Fig. 3. Implementation stage is the core of the whole emergency response, departments need to coordinate and cooperate to complete the rescue mission.

Emergency Recovery Stage Process

Emergency recovery stage mainly includes rehabilitation disposal, investigation, and evaluation. It is shown in Fig. 4. Through the recovery stage of the relevant disposal measures, the subway to resume normal operations, while the investigation of the accident assessment and summary, to provide reference for the future of the accident disposal.
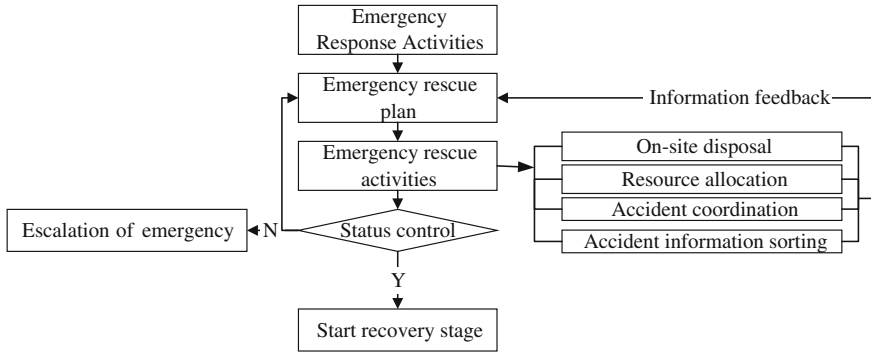
Fig. 3 Process of implementation stage
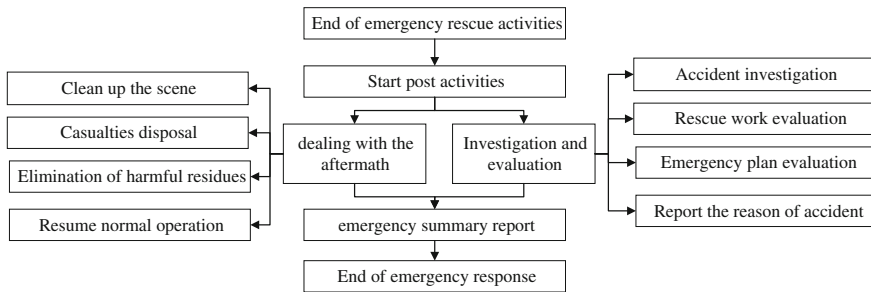
Fig. 4 Process of emergency recovery stage

Metro Emergency Response Process SPN Model

Modeling Basic Principle

According to the basic principle of Petri net, make full use of the advantage of Petri net in process modeling [12], stochastic Petri nets is used to modeling the emergency response process of Metro Emergency in this paper. In the SPN model, the transitions represent the emergency response activities; the places represent data processing and storage. The dynamic properties of the SPN mode can reflect logical order and the generation, storage, transmission, and processing of information and resources of the various activities in the process of emergency response well. Specific principles are embodied as follows:

(1) It is feasibility to model emergency process by using SPN to describe the process system, and the places represent information and other resources of the emergency process; the transition represents information production and transfer of the emergency process and describe the changes of information and other resources of emergency response process through assigning a delay time to the transition.

(2) It is enforceability of emergency process based on SPN. The place contains a token which indicates that the place is active and when the associated transitions are fired, the token is transferred from the place to another place which means that information, resources, and so on from one state to the next state in the process of emergency response. After the data tracking and collection activities are added in the model, the system can be simulated.

(3) It is evaluable of emergency process based on SPN. The response time and the operation efficiency of each link of the emergency response system can be evaluated by assigning the transition rate λ in the model.

The SPN Model for Each Stage of Metro Emergency Response

Through detailed analysis of each stage of the emergency disposal of Metro Emergency, we can set up the early warning, decision-making, response action, and recovery stages of the SPN model.

Early Warning Stage SPN Mode

The SPN model of early warning stage is shown in Fig. 5. A token placed in place P_1 when an emergency event occurs. The related personnel alarm is enabled expressed by transition T_1, when one placed in P_1. The firing transition T_1 consumes the one token in place P_1 and produces one token in place P_2, representing that Operating Control Center (OCC) received alarm information. Others places P_3–P_6, respectively, represent preliminary handing information, emergency classified information, application for higher-level response procedures, emergency decision scheme information, and emergency summary report. The transitions T_2–T_7,respectively, represent preliminary handing, cancel alarm, OCC determine emergency classification level, application for higher-level response procedures, report to the superior emergency leadership organization, and start rescue programs. The transitions T_1–T_7 rates are, respectively, represented by $\lambda_1 - \lambda_7$. The SPN model of early warning stage is shown in Fig. 5.

Decision-Making Stage SPN Mode

A token placed in place P_1 when the emergency classified level information is determined. Transitions T_1 and T_2 represent a conflict structure, and its firing is determined by token placed in place P_1. The firing transition T_1 consumes the one token in place P_1 and produces one token in place P_2, representing that emergency leadership organization received alarm information. The places P_3–P_{17},respectively,
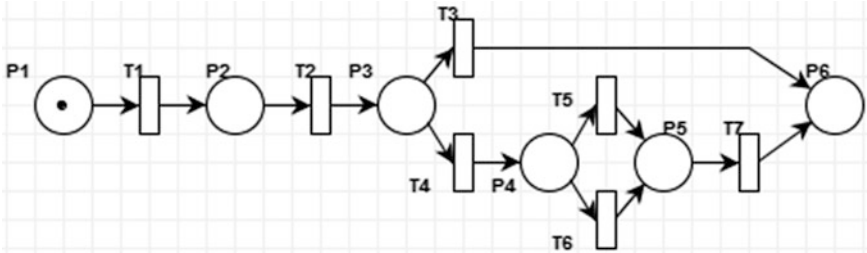
Fig. 5 Early warning stage SPN model
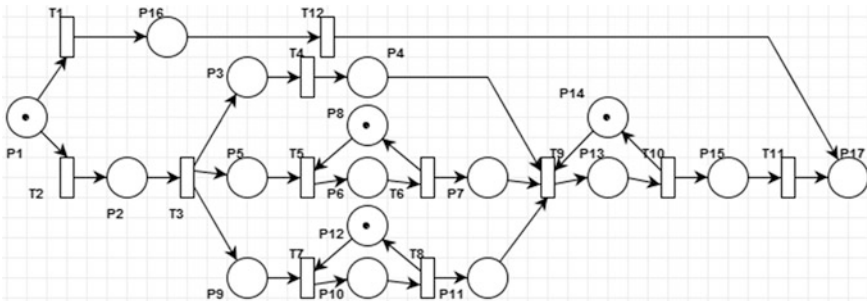
Fig. 6 Decision-making stage SPN model

represent emergency situation information, the result of emergency situation analysis, emergency plan information, preliminary emergency plan information, final emergency plan information, emergency plan revision information, expert database information, expert consultation information, final expert consultation information, expert consultation modify information, preliminary rescue plan, plan modify information, final rescue plan, and higher-level response information. The transitions T_3–T_{11}, respectively, represent that start alarm information analysis, on-site situation analysis, emergency plan analysis, call emergency plan, expert consultation, call expert consultation result, generate rescue plan, rescue plan evaluation, start rescue plan, and assist in higher-level response. The transitions T_1–T_{11} rates are, respectively, represented by λ_1–λ_{11}. The SPN model of decision-making stage is shown in Fig. 6.

Implementation Stage SPN Mode

A token placed in P1 when the rescue plan is determined. The firing transition T_1 consumes the one token in place P_1 and produces one token in place P_2, P_3, P_4, and P_5, respectively, represent that on-site disposal information, resource allocation information, accident coordination information, and accident information sorting information. The places P_6–P_{12}, respectively, represent that on-site disposal
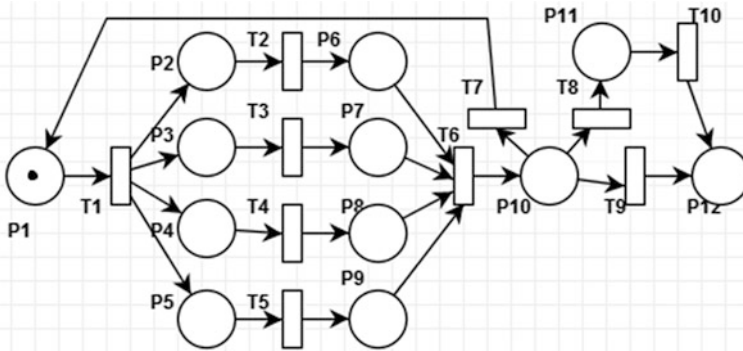
Fig. 7 Implementation stage SPN mode

feedback information, resource allocation feedback information, accident coordination feedback information, accident information sorting feedback information, situation analysis information, higher-level response information, and emergency summary report. The transitions T_2–T_{10}, respectively, represent that on-site disposal, resource allocation, accident coordination, accident information sorting, situation analysis, modifying rescue plan, escalation of emergency, start post activities, and assist in higher-level response. The transitions T_1–T_9 rates are, respectively, represented by λ_1–λ_9. The SPN model of decision-making stage is shown in Fig. 7.

Emergency Recovery Stage SPN Mode

There is not detailed to investigation and assessment of emergency process. A token placed in place P_1 when start post activities. The transition T_1 represents post-activities. The firing transition T_1 consumes the one token in place P_1 and produces one token in place P_2, P_3, respectively, represent that dealing with the aftermath information and investigation and evaluation information. The place P_4 represents emergency summary report. The transition T_2 represents emergency summary. The transitions T_1, T_2 rates are, respectively, represented by λ_1, λ_2. The SPN model of recovery stage is shown in Fig. 8.

Fig. 8 Emergency recovery stage SPN mode
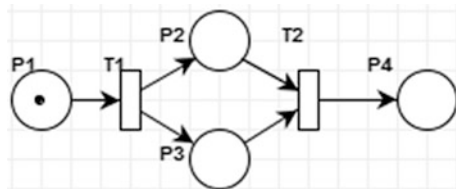
The SPN Mode of Metro Emergency Response

According to the preamble of metro emergency response, the contents of each step of process analysis and modeling, fully integrated Petri net theory of information disposal, with SPN model of each stage of public transitions and places, we establish a comprehensive subway emergency disposal of the whole process of Petri net model, which shown in Fig. 9. The definitions of the places P_k and transitions T_k are listed in Table 1 and Table 2.
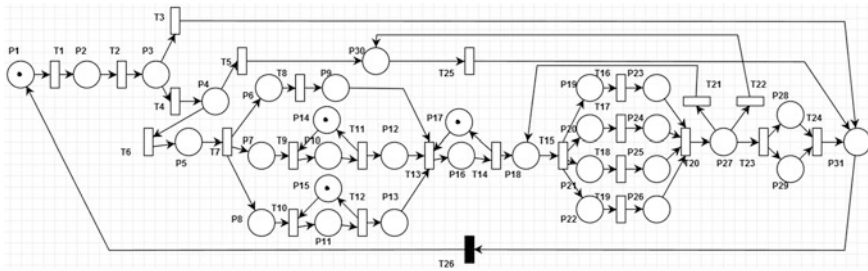
Fig. 9 Metro emergency response process SPN mode

Table 1 Definitions of places

Place	Definitions	Place	Definitions
P_1	Emergency message	P_{17}	Plan modify information
P_2	OCC received alarm information	P_{18}	Final rescue plan
P_3	Preliminary handing information	P_{19}	On-site disposal information
P_4	Emergency classified information	P_{20}	Resource allocation information
P_5	Emergency leadership organization received alarm information	P_{21}	Accident coordination information
P_6	Emergency situation information	P_{22}	Accident information sorting information
P_7	Emergency plan information	P_{23}	On-site disposal feedback information
P_8	Expert database information	P_{24}	Resource allocation feedback information
P_9	The result of emergency situation analysis	P_{25}	Accident coordination feedback information
P_{10}	Preliminary emergency plan information	P_{26}	Accident information sorting feedback information
P_{11}	Expert consultation information	P_{27}	Situation analysis information
P_{12}	Final emergency plan information	P_{28}	Dealing with the aftermath information
P_{13}	Final expert consultation information	P_{29}	Investigation and evaluation information
P_{14}	Emergency plan revision information	P_{30}	Higher-level response information
P_{15}	Expert consultation result modify information	P_{31}	Emergency summary report
P_{16}	Preliminary emergency plan information		

Table 2 Definitions of transitions

Transition	Definitions	Transition	Definitions
T_1	Related personnel alarm	T_{14}	Rescue plan evaluation
T_2	Preliminary handing	T_{15}	Start rescue plan
T_3	Cancel alarm	T_{16}	On-site disposal
T_4	OCC determine emergency classification level	T_{17}	Resource allocation
T_5	Application for higher-level response procedures	T_{18}	Accident coordination
T_6	Report to the superior emergency leadership organization	T_{19}	Accident information sorting
T_7	Start alarm information analysis	T_{20}	Situation analysis
T_8	On-site situation analysis	T_{21}	Modifying rescue plan
T_9	Emergency plan analysis	T_{22}	Escalation of emergency
T_{10}	Expert consulting	T_{23}	Start post activities
T_{11}	Call emergency plan	T_{24}	Emergency summary
T_{12}	Call expert consultation result	T_{25}	Assist in higher-level response
T_{13}	Generate rescue plan		

The instantaneous transition T_{26} is added in the input place P_{30} and the output place P_{30} to guarantee the continuous process of the network which the steady-state solution can be obtained, but count nothing actually. Transitions T_{3-4} and T_{5-6} represent a conflict structure. Transitions T_{8-10} and transitions T_{16-19} represent a simultaneous structure. There is loop structure in SPN model.

The SPN Mode Performance Analysis

Reachability Graph of SPN Mode

According to the SPN model shown in Fig. 9, from the initial state S_0 of the system, using the Petri net modeling tool, we can put all the state of the system and the relationship between the states. Reachability graph is shown in Fig. 10.

S_i represents the state of the system; S_4 represents the virtual state associated with the transition T_{25}. The reachability state of S_0 has 44. It can be determined that the model is live and bounded by reachable graph. Assign a relevant rate λ_i of transition Ti on each arc that can obtain isomorphic Markov chain. There are 43 states in reachability graph. Initial marking is denoted by $m_0 = (1, 13, 14, 16)$, representing the tokens that are these places P_1, P_{13}, P_{14}, P_{16} simultaneously. Assumed transitions T_1–T_{25} rates associated with correspondence parameter
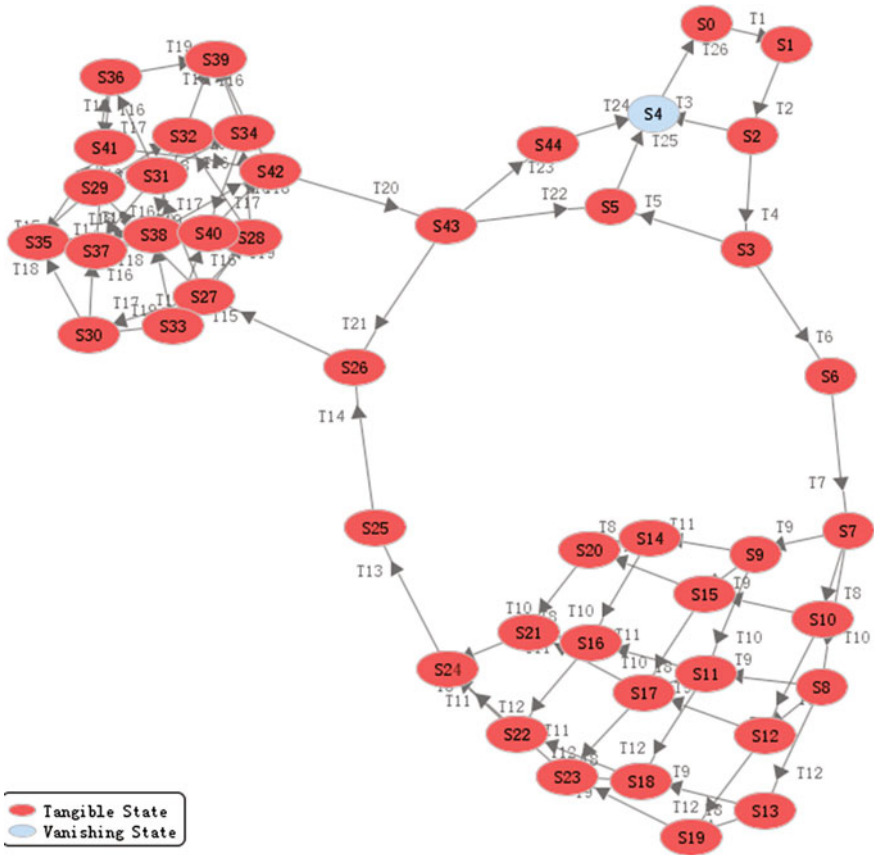
Fig. 10 Reachability graph

$\lambda_1 - \lambda_{25}$ and $x = (x_1, x_2, \ldots x_{44})$ are a set of steady-state probability of reachability states of Markov chain. We can obtain the steady-state probability based on correlation theorem of Markov chain [13], and calculate busy probability of places indicator and utilization indicator of transitions.

A Case Study of Beijing Metro

Any finite place, finite transition, marked stochastic Petri net are isomorphic to a one-dimensional discrete Markov process [14]. The performance of the average token number of places and the utilization of transitions are analyzed by steady-state marking probabilities of Markov chain.

   In this paper, a simulation analysis of Beijing Metro Emergency Plan response process is developed. This emergency plan divide emergency response process into

four stages such as early warning stage, decision-making stage, implementation stage, emergency recovery stage, and we can obtain SPN mode of this case based on metro emergency response process model in this paper. The Beijing metro emergency plan response process SPN model is shown in Fig. 11. The transitions of the decision-making stage is combined with the change of T14, and the transitions of emergency recover stage is merged into T24.

According to the experience of Beijing subway emergency experience data, we can assign each transitions T_i corresponding rate parameter λ_i which is shown in Table 3.

Solving the isomorphic Markov chain, we can obtain steady-state making probabilities shown in Table 4.

Calculate average token number of places and utilization of transition based on steady-state making probabilities shown in Tables 5 and 6.

(1) The average token number of places

It reflects the busy probability of the places. The average token number of places P23–P26 and P30 shown in Table 5 is significantly greater than other place, and in the process of emergency, information feedback and higher level of emergency response takes much longer compared to other links. Because in the process of emergency, emergency leading group of unified command departments launched rescue operations, which lead to multi departments rescue receiving feedback
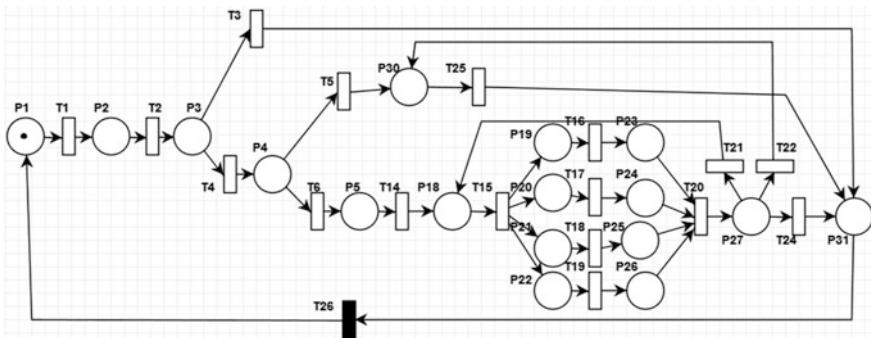
Fig. 11 Beijing metro emergency plan response process SPN model

Table 3 Transitions rate

Parameter	Rate value	Parameter	Rate value	Parameter	Rate value	Parameter	Rate value
λ_1	1	λ_6	1	λ_{11}	0.5	λ_{16}	1/12
λ_2	1	λ_7	1	λ_{12}	0.5	λ_{17}	1/24
λ_3	1	λ_8	1	λ_{13}	1		
λ_4	1	λ_9	0.5	λ_{14}	1		
λ_5	1	λ_{10}	0.5	λ_{15}	1		

Table 4 Steady-state making probabilities

Steady-state making probabilities	Value	Steady-state making probabilities	Value	Steady-state making probabilities	Value
P(M0)	0.04375	P(M8)	0.00467	P(M16)	0.00467
P(M1)	0.04375	P(M9)	0.00467	P(M17)	0.00467
P(M2)	0.02916	P(M10)	0.00467	P(M18)	0.01402
P(M3)	0.01458	P(M11)	0.00467	P(M19)	0.01402
P(M4)	0.01458	P(M12)	0.00467	P(M20)	0.01402
P(M5)	0.66777	P(M13)	0.00467	P(M21)	0.01402
P(M6)	0.02805	P(M14)	0.00467	P(M22)	0.02805
P(M7)	0.01402	P(M15)	0.00467	P(M23)	0.01346

Table 5 Average token number of place

Places	Average token number	Places	Average token number	Places	Average token number	Places	Average token number
P1	0.04375	P18	0.02805	P23	0.08881	P30	0.66777
P2	0.04375	P19	0.05609	P24	0.08881	P31	0
P3	0.02916	P20	0.05609	P25	0.08881		
P4	0.01458	P21	0.05609	P26	0.08881		
P5	0.01458	P22	0.05609	P27	0.01346		

Table 6 Utilization of transition

Transition	Utilization	Transition	Utilization	Transition	Utilization	Transition	Utilization
T1	0.04375	T6	0.01458	T18	0.02805	T24	0.00112
T2	0.04375	T14	0.01458	T19	0.02805	T25	0.02805
T3	0.01458	T15	0.02805	T20	0.02805		
T4	0.02916	T16	0.02805	T21	0.01346		
T5	0.01458	T17	0.02805	T22	0.01346		

information, and in a short period of time of information for decision-making, this case easily lead to information blockage, in addition to catastrophic events, the emergency response to the need for a cross organizational, cross area of city level emergency response, easy to produce a plethora of information, which leads to the accumulation of information. Therefore in the rescue process and need to pay attention to all kinds of information sorting, the improvement of emergency decision-making mechanism and other sectors of society to emergency response, strengthening of Metro emergencies emergency drills which is a key of optimizing the process of emergency.

(2) Utilization of transition

It reflects that the time is occupied by the activities in emergency response process. The utilization of Transition T1, T2, T4, T15–T20 from Table 6 is significantly higher than that of other changes high. In the process of emergency accident reporting, grade judgment, resource scheduling on-site, external coordination, the situation analysis activity for a longer period of time. Because the on-site disposal involves the multi-body, emergency command group coordination difficulty bigger, received too many accident information, decision required a longer period of time, in addition to meet emergency needs more resources, scheduling to step in place. Therefore, the emergency response organization should strengthen management, clear the duties of each department in emergencies, so as to enhance the various departments in the process of emergency coordination, enhance the efficiency of subway emergency disposal.

Conclusions

In this paper, a SPN model of metro emergency response process is constructed based on the stochastic Petri net. An isomorphic Markov chain constructed after the model is proved that it is active and bounded by constructing reachability graph. The average token number of places and the utilization of transitions are calculated based on Markov chain. According to the Beijing metro emergency plan response process, simulation analysis is carried out and we get busy probability of emergency response process of each link and the time of activities in emergency response process and we put forward suggestions to improve the emergency rescue efficiency. It plays a guided role to the information construction and improves the efficiency of multi department linkage of the metro emergency response.

References

1. Wang, Z.Q. 2015. Event driven simulation model research of urban rail transit emergency treatment process.*Application Research of Computers* (01): 125–131.
2. Zarboutis, N., and N. Marmaras. 2004. Searching efficient plans for emergency rescue through simulation: the case of a metro fire. *Cognition, Technology & Work* 6 (2): 117–126.
3. Zhang, Y.F., and W.Y. Chen. 2015. Markov chain analysis of metro network in responding to emergency. *China Safety Science Journal* 02: 165–170.
4. Sun, Q.Y., and X.Y. Li. 2015. Establishment of emergency coordination model across organizations based on stochastic Petri net. *China Safety Science Journal [J]* 09: 63–69.

5. Li, Q., Y. Deng, C. Liu, et al. 2016. Modeling and analysis of subway fire emergency response: An empirical study. *SAFETY SCI* 84: 171–180.
6. Liu, X., and Q. Xia. 2015. Research on the metro emergency management. *Technology & Economy in Areas of Communications* 01: 85–88.
7. Lu, W.Q. 2011. On constructing metro emergency management system: based on complex system theory perspective. (04):119–124.
8. Kaakai, F., Hayat, S., and El Moudni, A. 2007. A hybrid petri nets-based simulation model for evaluating the design of railway transit stations [J]. *Simulation Modelling Practice and Theory* 15 (8).
9. Landry, J.F., Ulmer, C., and L. Gomez. 2008. Fuzzy distributed workflows for crisis management decision makers. *Intelligence and Security Informatics* 5376.
10. Sultana, S, and Z. Chen. 2007. Modeling of flood-related interdependencies among critical infrastructures. Berlin: Spring.
11. Zhong, M., Shi, C., Fu, T., et al. 2010. Study in performance analysis of china urban emergency response system based on Petri net. *Safety Science* 48(6).
12. Lin, C. 2005. Stochastic Petri nets and system performance analysis. Beijing: Tsinghua University Press.
13. Jiang, Z.B. 2004. *Petri nets in manufacturing systems modeling and control*. Beijing: China Machine Press.
14. Mollog, M.K. 1982. Performance analysis using stochastic Petri nets. *IEEE Transactions on Computers* 31 (9): 913–917.

Layout Optimization Design of Electric Vehicle Charging Station Based on Urban Parking Lot

Yaqin He, Xinglin Zhou, Zupeng Liu and Maoping Ran

Abstract As a promising vehicle in the future, electric vehicles will become the common traffic modes for urban residents with the characteristics of green, environmentally friendly, and low carbon. The reasonable layout of the electric vehicle charging station is one of the key factors that will influence the promotion of electric vehicles. Taking urban parking as a platform and based on the harmonization with construction cost and user's interest, this paper established layout optimization model of charging station using transportation planning theory. Finally, an example is given to test the model. The results can provide certain decision guidance for government to plan electric vehicle charging station.

Keywords Urban traffic · Site layout · Transportation planning · Electric vehicle charging station · Parking lot

Introduction

With the increasing deterioration of global environment and the increasing tension of oil resources, the electric vehicle, as a new energy vehicle, was paid more attention. In order to promote the development of electric vehicles, improvement of the basic facilities is required. The layout of electric vehicle charging station is one of the important parts of electric vehicle facilities and takes very important role on pulling the development of entire electric vehicle industry.

At present, domestic and foreign scholars, experts have carried out a series of studies on electric vehicle charging station planning and achieved some results [1–13]. But the present results are achieved based on ideal conditions without considering the actual geographical condition. It does not consider whether the calculated location is proper to be site or not. Maybe it is in busiest commercial area or

Y. He (✉) · X. Zhou · Z. Liu · M. Ran
School of Automobile and Traffic Engineering, Wuhan University of Science
and Technology, Wuhan 430081, Hubei, China
e-mail: heyaqin@126.com

lakes. Therefore, with a combination of actual geographical factors and considering the shortage of urban land, this paper puts forward a new mode that electric vehicle charging station will be built based on urban parking lot, and discussed the charging station layout that has important theoretical significance and practical value for promotion and development of electric vehicles.

Layout Optimization of Charging Station

The Principle of Preliminary Address of Charging Station

The charging station is built using present parking lot can not only meet the user's demand and ensure feasibility of station, but also save the land and reduce the cost. Therefore, how to select the existing parking lot as the primary address reasonably is critical that can be conducted according to the following principles:

(1) The rebuilding parking lots are public parking lots;
(2) The number of berth of the rebuilding parking should be not less than 50 [14]
(3) The distance between the two rebuilding parking should be in accord with the service radius of charging station, $s_{\min} \leq \bar{s}_{ij} \leq s_{\max}$.

Considering the distance between the station i and the station j, then

$$\bar{s}_{ij} = \sqrt{(X_i - X_j)^2 + (Y_i - Y_j)^2}, i \neq j \tag{1}$$

where (X, Y) is the coordinates for charging stations.

In general, the distance between two stations cannot be too close that will cause the waste of resources. So, in order to meet the service radius, for any station, there is one station that is not far away and the distance between them is not more than a certain value, otherwise a large space cannot be service. Therefore, $\bar{s}_{ij}$ should meet the following equation $s_{\min} \leq \bar{s}_{ij} \leq s_{\max}$. According to Chen and Zhou [15], the minimum distance between two parking is 50 m, and considering the design range of the charging station, $s_{\min}$ will be set to 0.2 km. In order to assure that all the electric cars can be service in the area, we can take twice the service radius as $s_{\max}$. According to the service radius of gas station in the "Code for transport planning on urban road," the service radius is 1 km, namely $s_{\max}$ is 2 km.

Analysis of Influence Factors of Charging Station Distribution

There are many factors that can affect the layout of charging station, such as volume, service radius, environmental protection, regional distribution ability, road

network planning, investment cost, and operating expenses. Considering various factors, this paper tries to take into account investors and users, namely the investors will provide the best service using the least economic cost. It focuses on the following three aspects:

(1) The social cost of utilization of electric vehicle charging station, here it is travel cost. Every user hope to receive service at the nearest charging station, so the longer the travel distance the greater social cost;
(2) The investment cost of building electric vehicle charging station. Although the electric car industry is strongly supported by the state, investors still hope that the minimum investment can provide the best service.
(3) Rebuilding electric vehicle charging station will inevitably reduce the area of parking that will bring a certain loss. It is expressed by economic loss of parking lot.

Optimization Model for the Location

User's Social Costs C_1

According to transportation planning theories, the area used to set electric vehicle charging station is divided into several zone, as shown in Fig. 1. The division principle is as follows [16, 17]:

(1) It can be divided according to administrative divisions;
(2) The service radius of electric vehicle charging station should be considered;
(3) The density of electric cars should be considered.

Firstly, the area is divided into n small zones. It is assumed that each zone has an electric vehicle production point (the center of the electric vehicle charging demand), then, the production point set of electric vehicle charging demand can be expressed by $E = \{E_1, E_2, \ldots E_n\}$, where E_i is the ith electric vehicle production point, and its coordinate is (x_i, y_i). The charging demand of the ith electric vehicle production point is d_i. The set of the alternative location of charging station

Fig. 1 Map of district division
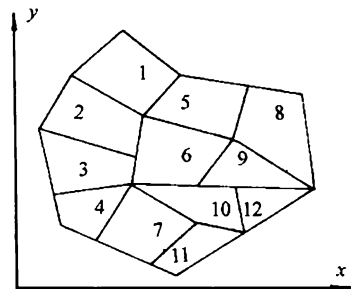
(parking lot) is $S = \{S_1, S_2, \ldots S_m\}$, where, S_j is the jth charging station, and its coordinate is (X_j, Y_j). P_j is service capacity provided by the jth charging station. The service capacity of charging station in the zone should meet the electric vehicle charging demand as possible, but it is limited with parking lot condition. Therefore, there has $P_j^{min} \leq P_j \leq P_j^{max}$.

Assuming that A_{ij} is the number of electric vehicles that are produced in zone i and accept service in the charging station j. Then the following equations are tenable,

$$\sum_{i=1}^{n} A_{ij} = P_j, \sum_{j=1}^{m} A_{ij} = d_i \text{ and } A_{ij} \geq 0.$$

l_{ij} is called charging distance that is defined as the distance between the electric vehicle production point i and the charging station j and can be expressed as follows:

$$l_{ij} = \sqrt{(x_i - X_j)^2 + (y_i - Y_j)^2} \tag{2}$$

So, the total traveling distance is as follows:

$$T = \sum_{i=1}^{n} \sum_{j=1}^{m} l_{ij} \cdot A_{ij} \tag{3}$$

The limited distance between the generating point and charging station can be expressed by D that is the distance that drivers can accept the maximum traveling distance. Through the questionnaire on the drivers, the tolerable time taken to search for a charging station can be obtained, and it is 10 min. The travel speed on urban road is generally 40 km/h, so D is taken as 7 km. a_{ij} was defined as the variable that describes how to select charging station by drivers, there is

$$a_{ij} = \begin{cases} 1, & l_{ij} \leq D \\ 0, & l_{ij} > D \end{cases} \tag{4}$$

$a_{ij} = 1$ indicates the charging station j can provide service for the generating point i. $a_{ij} = 0$ indicates the charging station j cannot provide service for the generating point i.

Assuming that the every generating point in the region has corresponding charge station that can provide service, and each charging station must have the corresponding electric vehicle as the service object, there are as follows:

$$\sum_{i=1}^{n} a_{ij} \geq 1, \sum_{j=1}^{m} a_{ij} \geq 1 \tag{5}$$

All the users hope to charge at the nearest charging station when the electric car has no power, namely the traveling distance is shortest that indicates the user's social cost is least. Assuming the unit distance cost is expressed by R, the total cost for the users is as follows:

$$C_1 = R \sum_{i=1}^{n} \sum_{j=1}^{m} l_{ij} \cdot A_{ij} \tag{6}$$

The Investment Cost of Building Additional Charge Station C_2

The investment cost of charging station includes such fixed cost as storage battery and other facilities and construction cost of charging pile. It can be expressed as

$$C_2 = F + G = \sum_{j=1}^{m} (F_j b_j + G_j b_j) \tag{8}$$

b_j was defined as the decision variables, there is

$$b_j = \begin{cases} 1 \\ 0 \end{cases} \tag{6}$$

$b_j = 1$ denotes that charging station j is selected, $b_j = 0$ denotes that the charging station j is not selected. where F_j represents the fixed cost of charge station j that can be considered as a constant.

G_j represents the construction cost of charge pile in charging station j, the calculation formula is as follows:

$$G_j = P_{jk} Q_{jk} \tag{9}$$

P_{jk} is the number of the charging pile with k power in the charging station j. The total number of charging pile is P_j that is the service capacity of charging station j.

Assuming that the area of charging station (parking lot) can be expressed by $B = \{B_1, B_2, \ldots B_m\}$, where B_j denotes the area of parking lot j.

In order to ensure parking supply is not reduced after construction additional charging stations, a three-dimensional park can be built. The area of three-dimensional park is expressed by $C = \{C_1, C_2, \ldots C_m\}$, where C_j denotes the area of three-dimensional park in parking lot j.

After the reconstruction of the parking lot, the total area of aisle and lane is $U = \{U_1, U_2, \ldots U_m\}$, where U_j denotes area of aisle and lane of parking lot j.

E is the area of each berth according to criteria.

The number of surface berths after reconstruction can be expressed by $H = \{H_1, H_2, \ldots H_m\}$, H_j denotes the number of surface berths in the parking lot j, so

$$H_j = \left(B_j - C_j - U_j\right)/E \tag{10}$$

Considering that charging piles are only installed on surface berths, so there is

$$P_j \le \left(B_j - C_j - U_j\right)/E \tag{11}$$

The price of charging pile is expressed by $Q = \{Q_1, Q_2, \ldots Q_s\}$ Q_k denotes price of charging pile with power K.

$$C_2 = \sum_{j=1}^{m} \left(F_j b_j + G_j b_j\right) = \sum_{j=1}^{m} F_j b_j + \sum_{j=1}^{m} \sum_{k=1}^{s} P_{jk} Q_{jk} b_j \tag{12}$$

The Economic Loss of Parking Lot C_3

In order to offset the economic loss resulted from reduction of parking lot area, a three-dimensional park is built to guarantee parking supply. So the economic losses of parking lot can be expressed by cost of building a three-dimensional park C_3.

$$C_3 = \sum_{j=1}^{m} O_j b_j \tag{13}$$

where O_j denotes the cost of building a three-dimensional park in parking lot j.

Site Layout Optimization Model

Based on the minimum economic cost and better service for customers, the site layout model of charging station is built

The objective function:

$$
\begin{aligned}
\min(C) &= \min(C_1 + C_2 + C_3) \\
&= \min\left(R \sum_{i=1}^{n} \sum_{j=1}^{m} l_{ij} \cdot A_{ij} + \sum_{j=1}^{m} F_j b_j + \sum_{j=1}^{m} \sum_{k=1}^{s} P_{jk} Q_{jk} b_j + \sum_{j=1}^{m} O_j b_j \right)
\end{aligned}
$$

Constraint conditions:

$$
\begin{cases}
\displaystyle\sum_{i=1}^{n} A_{ij} = P_j \\[2mm]
\displaystyle\sum_{j=1}^{m} A_{ij} = d_i \\[2mm]
A_{ij} \geq 0 \\[2mm]
a_{ij} = \begin{cases} 1, & l_{ij} \leq D \\ 0, & l_{ij} > D \end{cases} \\[2mm]
\displaystyle\sum_{i=1}^{n} a_{ij} \geq 1 \\[2mm]
\displaystyle\sum_{j=1}^{m} a_{ij} \geq 1 \\[2mm]
t_{ij} = \sqrt{(x_i - X_j)^2 + (y_i - Y_j)^2} \\[2mm]
P_j^{\min} \leq P_j \leq P_j^{\max} \\[2mm]
P_j \leq (B_j - C_j - U_j)/E
\end{cases}
$$

where d_i, $P_j^{\min}$, $P_j^{\max}$, F_j, O_j, U_j, B_j, C_j, and E are the known quantities; A_{ij} and P_j are the unknown quantities.

Case Analysis

Assume that a region is divided into three zones, $E = \{E_1, E_2, E_3\}$. According to the layout principle of charging station, four parking lots are selected as preliminary site of electric vehicle charging stations, $S = \{S_1, S_2, S_3, S_4\}$. The charging production of three zones can be expressed by $d = \{45, 75, 60\}$, and the maximum service capacity and minimum service capacity of four selected charging stations is can be expressed by $P_{\max} = \{75, 85, 65, 100\}$ and $P_{\min} = \{35, 45, 25, 50\}$, respectively. The distance between three electric car demand points and the four stations can be expressed by the matrix $l_{ij} = \begin{pmatrix} 3 & 4 & 3 & 5 \\ 5 & 6 & 5 & 4 \\ 3 & 2 & 3 & 6 \end{pmatrix}$. In order to simplify calculations, the investment cost of reconstruction charging station C_2 and economic loss of parking lot C_3 were assumed to be permanent, so here only user social costs C_1 is considered.

Based on the above established model, this question can be solved by the minimum element method of transportation problem in operational research [18]. The final optimal solution is that S_1, S_2, S_3 are taken as electric car charging stations.

Conclusions

To speed up the promotion and popularization of electric vehicles has a great significance for energy conservation, emissions reduction, and energy structure upgrade. Charging station is one of the important bases for the development of electric vehicle. The rationality of the layout of charging station has a great influence on popularization and operation the electric car. In this paper, the layout optimization of charging station was discussed based on urban parking lot. And an example was taken to test the model. But not all the influenced factors were considered when establishing model, and some parameters in the model should be further identified, the model should be also improved in the application.

References

1. Owen, S.H. 1998. DASKIN M S. Strategic facility location: A review. *European Journal of Operational Research* 111 (3): 423–447.
2. Andy Ip, Simon Fong, and Elaine, Liu. 2010. Optimization for allocating BEV recharging stations in urban areas by using hierarchical clustering. In *The 2nd International Conference on Data Mining and Intelligent Information Technology Applications (ICMIA 2010),* Seoul, Korea.
3. Morrow, K., D. Karner, and J. Francfort. 2008. *Plug-in hybrid electric vehicle charging infrastructure review*. Idaho Falls: Idaho National Laboratory.
4. Revelle, C.S., and H.A. Elselt. 2005. Location analysis: A synthesis and survey. *European Journal of Operational Research* 165 (1): 1–19.
5. Zhang Changhua, and XIA Aihua. 2010. A novel approach for the layout of electric vehicle charging station. Chengdu: Apperceiving Computing and Intelligence Analysis.
6. Zhang Guoliang, Li Bo, and Wang Yunfa. 2011. Location and algorithm of multi-level electric vehicle charging stations. *Journal of Shandong University (Engineering Science)* 41(6):136–142.
7. Shaoyun, G., F. Liang, L. Hong, and W. Long. 2013. Planning of charging stations considering traffic flow and capacity constraints of distribution network. *Power System Technology* 37 (3): 582–588.
8. Zhou Hongchao, and Li Haifeng. 2011. Optimization model of electric vehicle charging station siting based on game theory. *Science Technology and Industry* 11(2):51–54.
9. Lingfeng, K., L. Zifa, and Z. Huan. 2011. Modeling algorithm of charging station planning for regional electric vehicle. *Modern Electric Power* 27 (4): 44–48.
10. Ren Yulong, Shi Lefeng, and Zhang Qian. 2011. Optimal distribution and scale of charging stations for electric vehicles. *Automation of Electric Power Systems* 35(14):53–57.
11. Liu Zhipeng, Wen Fushuan, and Xue Yusheng. 2012. Optimal Siting and Sizing of Electric Vehicle Charging Stations. *Automation of Electric Power Systems* 36(3):54–59.
12. He, Zhanyong. 2012. Research on planning method and operation mode of electric vehicle charging station. Beijing Jiaotong University.
13. Wang, Hui. 2013. Research on planning and operation of electric vehicle charging. Zhejiang University.

14. Ministry of Construction of the People's Republic of China. Code for transport planning on urban road (GB 50220-95). 1995-1-14.
15. Chen Jun, and Zhou Zhiyong. 2007. Urban parking facilities planning methods and information guidance technology. Southeast University Press.
16. Xu Weici. 2000. Urban transportation planning theory. Tongji University Press.
17. Hu Liege, and Liu Zhong. 2009. Transport junction and port station. China Communications Press.
18. Operations Research (3rd ed), Tsinghua University Press. 2012.

Study on the Changes of EEG Signal and Driving Behavior Based on the Driving Simulator

Guangyin Han and Liu Yang

Abstract In this paper, five participants take part in the driving simulator experiment. The simulated urban road is not interrupted, and the drivers' electroencephalogram (EEG) data, driving simulator data, and experimental video data are recorded synchronously. We put vehicles that change lanes randomly in the experimental scene to simulate the real situation, thus to get the changes in brain activity and the driving behavior. The power spectrum density and the energy of each frequency band of EEG signal are extracted by using the fast Fourier transform (FFT). The result indicates that (1) the beta wave has obvious advantages, when the driver faces a suddenly lane change in the front vehicle during continuous driving; the driver focuses on the front vehicle, and responds quickly to the events. (2) the female drivers are generally more stable than male drivers, but the response to environmental change is slower than that of the male drivers. (3) the drivers' first reaction when facing the suddenly lane change is to press the brake pedal and then steer off.

Keywords Driving simulator · EEG · Driver behavior · Power spectrum analysis

Introduction

The brain is the material base of all human advanced behavior, consisting of 10–160 billion nerve cells and 100 trillion synapses, and ten times more than that of the glial cells. The surface of the brain is divided into frontal, parietal, occipital, and temporal lobe by the central sulcus, occipital fissure, and lateral cerebral fissure. Among them, the temporal lobe is mainly for auditory function, the occipital lobe is

G. Han · L. Yang (✉)
MOK Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China
e-mail: dr.yangliu@hotmail.com

G. Han
e-mail: 14120773@bjtu.edu.cn

mainly for visual function, the parietal lobe is the senior center of body sensation, and the frontal lobe is mainly for the motor function of the body [1].

The brain waves are divided into four bands based on the difference of the frequency of the brain waves: delta (0.5–4 Hz), theta (4–8 Hz), alpha (8–13 Hz) and beta wave (13–30 Hz) [2]. Different people have different brain rhythms in various situations. In general, there is a dominant frequency which is the most prominent and obvious rhythm in the record [3]. It is most widely accepted that the function of electroencephalography (EEG) in different frequency bands is remarkably different at present [4]. The delta waves can be found in the deep sleep state. The theta waves can be recorded in adult sleepy state, indicating that the performance for the central nervous is inhibited. The alpha waves indicate that the brain is awake and relaxed, and it is easy to concentrate on study. People are awake when beta wave is the dominant wave and they are nervous and sensitive to the surroundings. EEG signal is mainly used in auxiliary diagnosis and treatment of mental disorders, such as epilepsy and schizophrenia [5], but fewer are mentioned in driving behavior evaluation studies. Drivers will fatigue with long time driving and be slower in reacting to the changes of the environment [6].

Fast Fourier transform (FFT) is a common method to extract features of the EEG [7, 8]. The EEG data are enormous with a long time of experiment, so FFT is a simple and fast calculation method. In this paper, the FFT is used to calculate the energy and power spectrum density of each wave band. The changing process is analyzed to research the driver's brain waves and driving behavior in the condition of the sudden change of the vehicle.

Materials and Methods

Experimental Equipment

The experiment uses a driving simulator that comes from the Beijing Jiaotong University (BJTU Simulator). The vehicle has the same size with Ford real car. The BJTU simulator can simulate a variety of traffic environments and driving scenes. It can also carry out relevant experimental operation, behavior data acquisition, and data integration.

Experiment Participants

Five healthy right-handed participants (aged 24–55 years old, two males, three females, all of them have more than three years of driving experience) take part in this study. Both professional drivers and nonprofessional drivers are consisted in this experiment. Here are the numbers of experiment participants: No. 1 (nonprofessional

male driver), No. 2 (nonprofessional male driver), No. 3 (professional male driver), No. 4 (professional female driver), and No. 5 (nonprofessional female driver).

Experiment Design

The experiment design referred to Ref. [9]. Experiment scene is consisted of urban ring road with two-way four lanes (the length of the road is 22 km), free flow of traffic with no pedestrians, good weather, and the speed limit is 80 km/h. All the experiments are conducted at the same time (10 a.m.–11 a.m.), and the participants do not know the experiment scene before the experiment. Experimental environment is stable and quiet. The vehicle which added in the experiment could change the lane randomly, and it may be in front of master of the car or in the opposite lane. The characteristics of these vehicles are of a suddenly lane change, and all of them are randomly, thus to simulate the condition that the driver encountering vehicle emergency lane change in the real life.

The participants should take a 10 min' adaptive training (not the experimental scene) before the experiment to test if the participants had no adverse reaction. The experiment is conducted in the condition of no other influence. The experiment is performed in about 20 min, and participants complete the starting, acceleration, steering, and lane changing, according to the road condition. The changes of driving behavior, brain waves, and driving simulator are recorded synchronously.

EEG Data Collection and Preprocessing

Neuroscan system is used to collect EEG data. Using the 64 lead measurements, the electrodes are placed in accordance with the international 10/20 system. The sampling rate is 1000 Hz, and the lead impedance is less than 10 KΩ.

EEG has been proved to be a typical nonlinear and nonstationary signal of low dimensional chaotic time series [10, 11], but the EEG signals can be regarded as stationary signal which was divided into short time series. The EEG data are preprocessed by removing eye movements and false trace movement [12, 13]. EEG data are off-line corrected form band pass filter, baseline correction, and independent component analysis (ICA) is adopted to remove artifacts [14].

EEG Energy of Each Frequency Band

Fast Fourier transform is one of the most commonly useful methods in EEG signal processing. It is a kind of method to transform the signal from time to frequency. In this paper, the time domain signal is transformed into the frequency domain signal by

Fourier transform, and the signal is reserved in the relevant frequency band, and then the characteristic wave is extracted by Fourier inverse transform. Signal frequency domain analysis is based on the Fourier transform, which transforms the time domain signal into frequency domain signal, thus to help people to understand the characteristics of the signal. Fourier transform formula is as shown in the Formula (1):

$$F(jw) = \int\limits_{-\infty}^{+\infty} f(t)e^{jwt}dt$$

$$f(t) = \frac{1}{2\pi} \int\limits_{-\infty}^{+\infty} F(w)e^{jwt}dw$$

(1)

We use the rhythm of the band energy of signal characteristic recognition as a significant feature to make the analysis more quantitative and intuitive [15], namely a finite discrete signal energy can be calculated by Formula (2).

$$E = \sum_{n=1}^{N} |x(n)|^2$$

(2)

which is the amplitude of the discrete points in the time domain, N for discrete points. According to Formula (2), we can calculate the absolute energy of each section of the delta, theta, alpha, and beta bands.

Results

EEG Data Analysis and Discussion

After the preprocessing, brain waves are as shown in Fig. 1, which stand for a male professional driver after pretreatment part of the electrode. Horizontal coordinate represents time (19–24 s, a total of 5 s). Vertical coordinate represents the name of electrode.

EEG data labeled all data segments on average superposition, firstly, after pretreatment, and then the power spectrum density and energy distribution are as shown in Fig. 2a, b.

The power spectrum density changes of delta, theta, alpha, and beta are as shown in Fig. 2a, b represent the energy distribution of 64 electrodes in beta, alpha, delta, and theta bands. The change of color represents different electrodes. According to the experimental results, FFT is accurate to extract the four basic rhythms, which are useful for the analysis of the energy distribution of the situation. By comparing the participants, we can find that the beta wave has the highest power. The experimental result shows that the beta wave is the dominant wave during driving simulation, and it means that the driver is conscious and the brain activity of whom
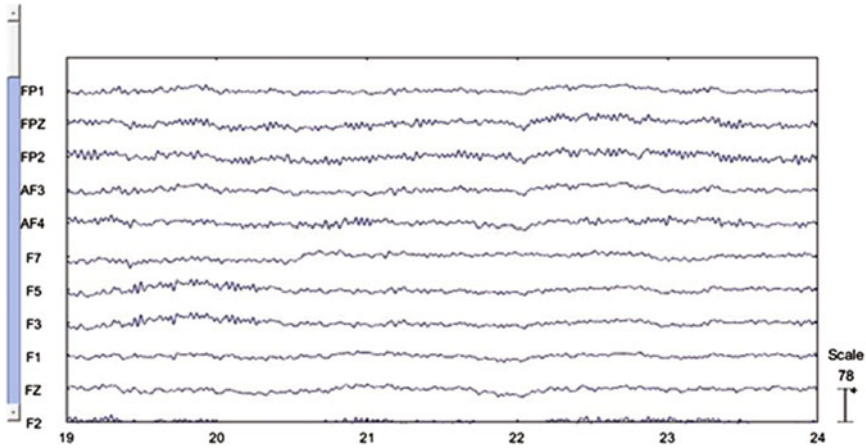
Fig. 1 Time domain map of EEG after pretreatment

is very active. The participants pay attention to surroundings, as well as they are also sensitive to the change in environment. At the same time, a long time driving will lead to driver fatigue and responses slowly to environmental changes.

Take frequency 6.0, 10.0, and 22.0 Hz which represent the theta, alpha, and beta waves as examples to draw the scalp power distribution diagram (delta wave is in the deep sleep state, thus it is out of consideration). The horizontal axis represents the frequency, while the vertical axis represents power, as shown in Fig. 3. It is the scalp power distribution of No. 3 participant, when he encountered the preceding vehicle changed lane. It can be seen that the beta wave in the frontal region has higher power, indicating that the participant's body movement was very active. Combined with video and driving simulator data, it can be found that when the participant suddenly encounters lane change in the front vehicle, the participant will steer, and manages to avoid the vehicle collision.

The EEG analysis of No. 3 participant when the front vehicle lateral lane changes is shown in Fig. 4. The electrode F4, which is in the frontal region, is selected for analysis because it has close relationship with the human body movement. The frequency bands range from 8 to 30 Hz, and the significance level is 0.01. The first subgraph is the ERP image, above the first subgraph, there is the electrode location, which the spot represents electrode position. The color change from blue to red in the phase-sorted trials represents power from low to high. It can be seen that the power reaches the maximum in the occurrence of a front vehicle steering. It indicates that the participant focuses on change in the surroundings. The second subgraph shows the changes of EEG amplitude during stimulation in F4 electrode. The third subgraph is ERSP (event-related spectral power), which represents event-related power mean changes (unit is dB). The curve does not exceed the blue region that means in the selected frequency 9.277 Hz relative to the baseline level 6.895 dB, the power does not have significant changes. The last subgraph is ITC graph (Inter-Trial coherence), which represents the degree of phase
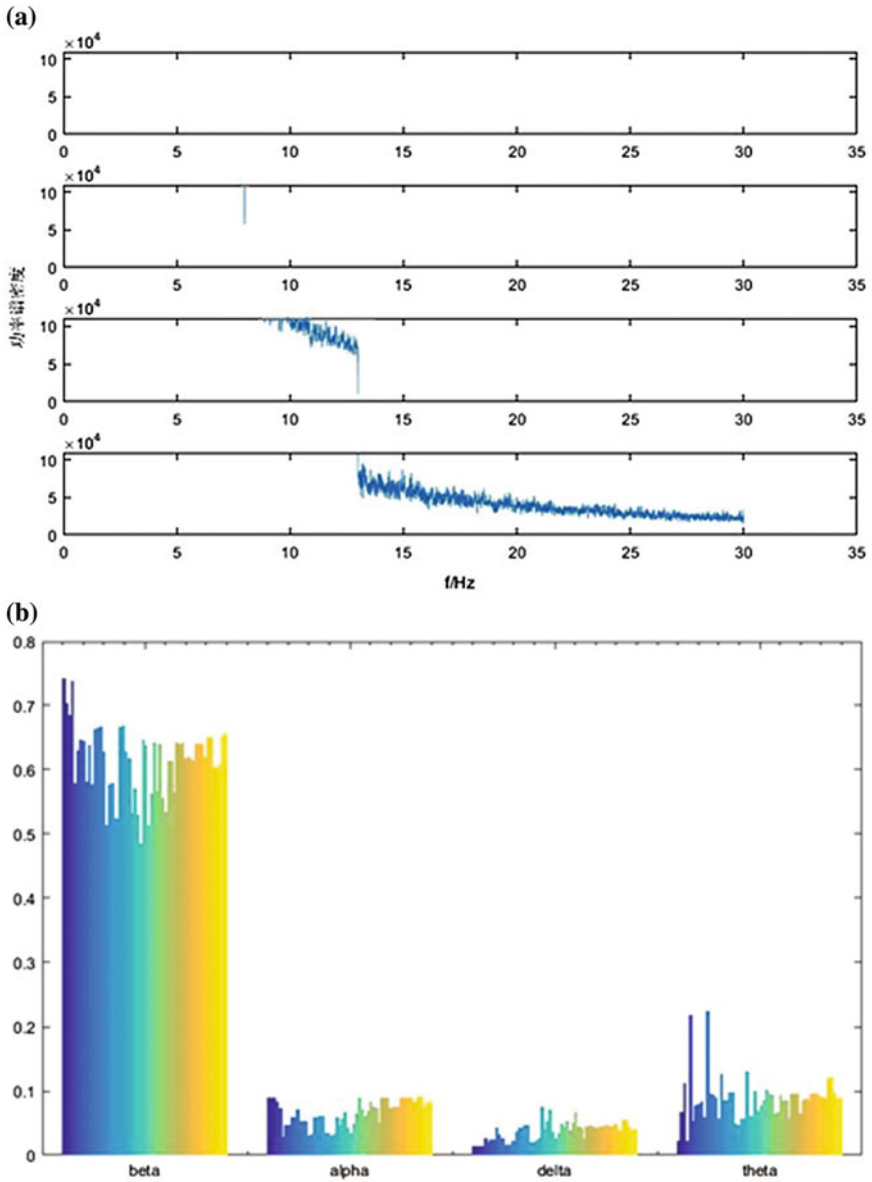
Fig. 2 a Power spectrum density distribution in each band **b** Energy distribution of each electrode in each band

synchrony of relative stimulation (driver's car encountered the preceding vehicle's suddenly lane change). The frequency of 9.277 Hz represents the selected band. It can be seen that the synchronization is significantly enhanced in the vicinity of the 1200 ms.
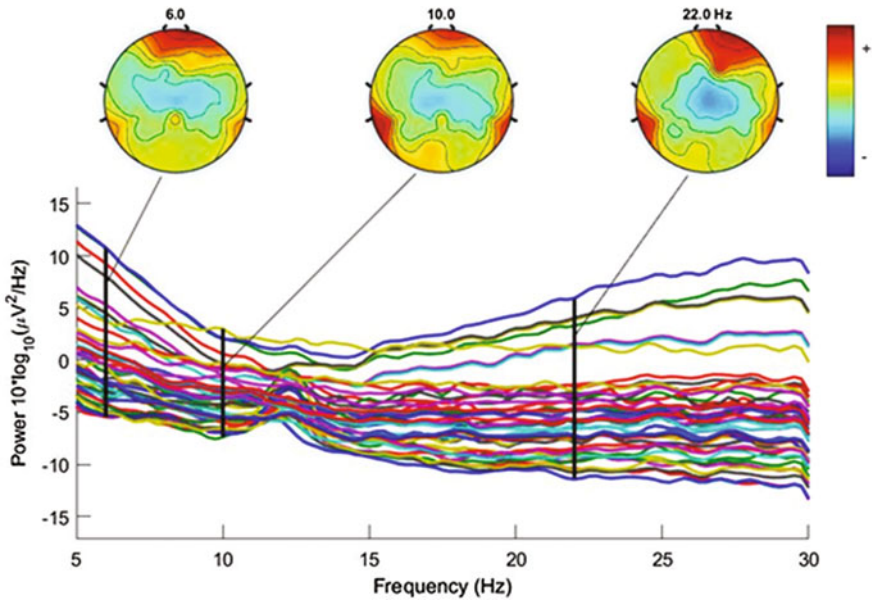
Fig. 3 Electrode spectra and maps
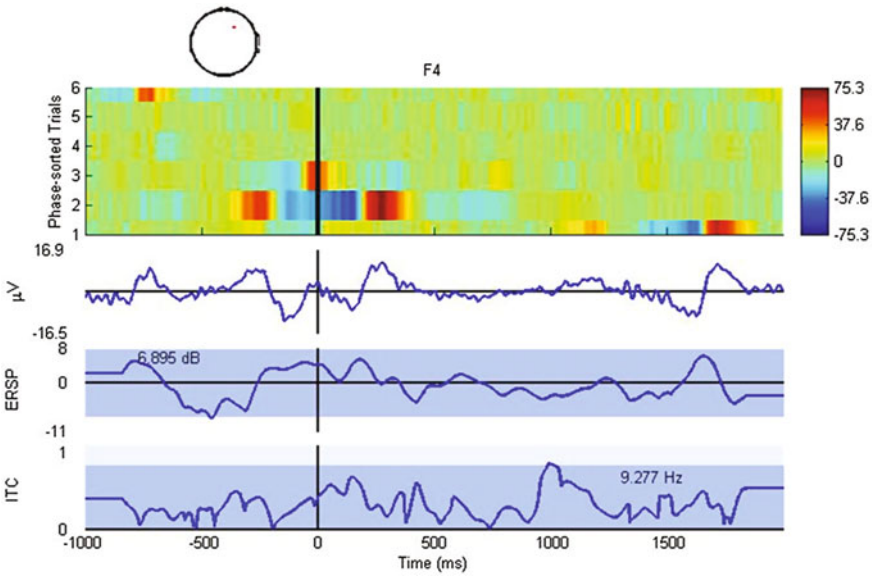
Fig. 4 ERP image, ERSP and ITC

Fig. 5 Voltage change of the electrode
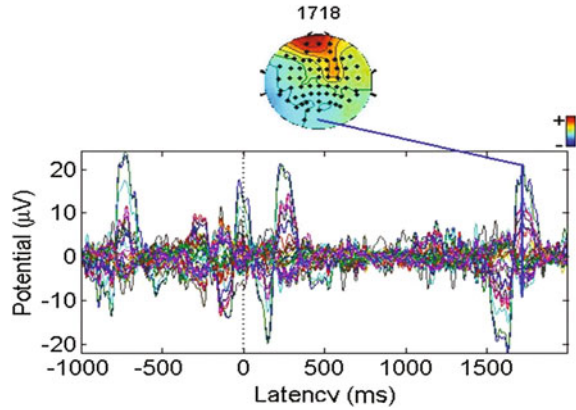
Event-related potentials (ERP) is a kind of brain image of cognitive science. The potentials change has close relationship with human physical or psychological activity, which relates to the electrical activity of the brain. The ERP is separated from EEG, which is recorded on the surface of the scalp after signal filtering and superposition. All of the five participants' ERP data are averaged on a single axis with scalp maps, as shown in Fig. 5. Each curve corresponds to each electrode, and the average voltage of the topography above is at 1718 ms (ERP maximum variance time map). It represents each electrode voltage distribution topographic map, when the front vehicle change lanes. It indicates that the voltage change reached the maximum at 1718 ms. According to the observation of the video and the driving simulator data, it can be found that when the front vehicle change lanes suddenly, the following driver will take a brake, and if the space is too narrow between them, the driver will not only take the emergency brake but also steer to avoid a collision.

Driving Simulator Data Analysis and Discussion

According to the video data, the driving operations of each participant when their front vehicle change the lane suddenly are obtained and statistical analysis, as shown in Table 1.

Table 1 Event statistics for participants

Participant number	Total number of the front vehicle steering	Total number of collision	Total number of deceleration	Total number of drivers' steering	Total number of drivers' deceleration and steering
1	8	1	4	1	2
2	10	0	6	2	2
3	11	0	5	2	4
4	10	0	6	1	3
5	12	1	6	2	3

(a)

(b)
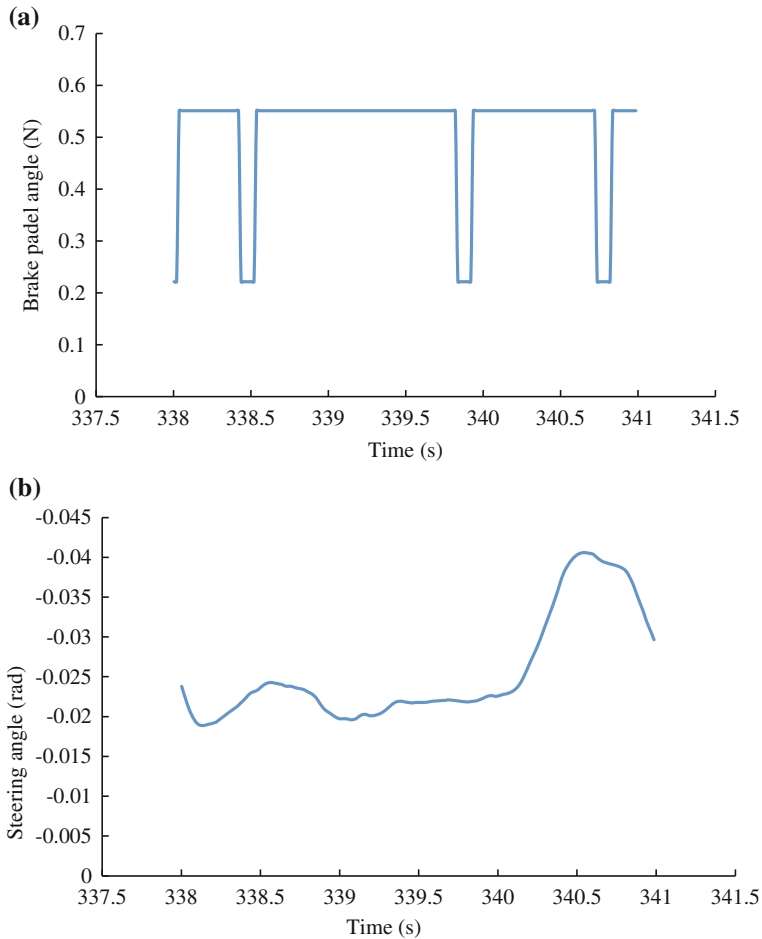
Fig. 6 **a** Change of brake pedal angle. **b** The change of steering angle

Five participants have suffered a total of 51 times of the front vehicle lane change, almost 10.2 times per person. When the drivers suffer sudden lane change of the front vehicle, they usually take three kinds of measures, that are deceleration, steering off, or both deceleration and steering. The participants who decelerate are 27 times in total, accounting for 52.94%. The proportion of the participant who chooses to steer account for 23.53%, and the proportion of the participant who slow down and steer off synchronously account for 23.53%. The result shows that the first reaction of both the male drivers and the female drivers in the condition of the front vehicle's lateral direction change is to step on the brake, and if the situation is very emergency, then the driver will start steering off to avoid collision.

There are two participants that have collision with other vehicles. In order to analyze the whole process of collision, No. 1 participant is selected as an example.
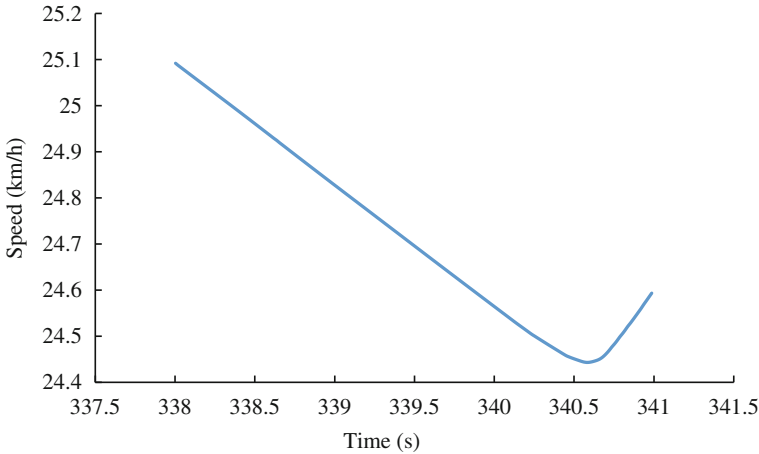
Fig. 7 Change of driver's vehicle speed

The change of the angle of the brake pedal is shown in Fig. 6a, and the change of the steering wheel angle during the collision of No.1 participant is shown in Fig. 6b. The horizontal coordinate represents the time of the collision. As shown in Fig. 6a, the participant takes three brakes when the front vehicle lane change. Combined with Fig. 6b, we can see that the driver does not steer off until the third brake. It indicates that collision can not be avoided after the second brake; thus, the participant take brake and steering control strategy almost at the same time.

The speed of No. 1 participant during the collision is as shown in Fig. 7. Combined with Fig. 6a, b, we can see that the speed decreases continuously, and the minimum value is reached at about 340 s, which is the time after the second brake. Furthermore, we can see that the speed starts to increase after the bottom point; the reason is that the participant can continue driving after the collision in simulated driving condition.

Conclusions

According to the analysis of EEG data and driving simulator data, we can draw the following conclusions:

(1) The beta wave is the dominant wave under the normal driving condition, and the drivers are sensitive to the surroundings. It indicates that the human brain is conscious, but also in a tension state. However, the driver's reaction to environmental changes will be slow with long time of continuous driving. According to the EEG band energy distribution map, we can see that the participant is consciousness as well as nervous.

(2) The professional drivers are found to be more stable compared to the non-profession drivers that have a smaller change of speed. The male driver reacts faster than the female drivers when the front vehicle lane change happened, while the female drivers are more stable while driving. It indicates that the professional drivers in the experiment do not have vehicle collision, and mostly they usually choose to brake and steer when the front vehicle changes the lane.

(3) EEG data combined with driving simulator data and video data can be fully synchronously used for analysis of the EEG changes and driving behavior changes which also can analyze the brain activity of various driving behaviors. EEGLAB is a good method to analyze EEG data. The interdisciplinary integration of neuroscience and traffic engineering, and using mathematical science methods and tools to analyze data, can improve the understanding of the relationship between the brain activity and the driving behavior.

Although the driving simulator can simulate the real road, the driver's mental state on the real road still has some differences. The future work is to collect EEG signal on real road and observe the changes of driver behaviors. Comparing EEG signals and driving behaviors in real and simulated situations can help research on the occurrence mechanism of hazardous driving behaviors, related traffic accident, and corresponding avoidance measures, which is meaningful to mitigate accident rate and improve drivers' ability to avoid accident.

Ethics Statement: The research involving human participants in this study has been approved by the Beijing Jiaotong University's research committee (per IRB). The written informed consent form for the experiment was also signed by each participant in this study.

References

1. Wang, Jin hui, Liang, Xia, and He Yong. 2013. Research on the human brain connecting group: brain structure network and functional network. *Science Bulletin* 2010 (16): 1565–1583.
2. Yang, Ren huan, Song, Ai guo, and Xu Bao guo. 2008. Analysis of basic rhythm of brain waves based on harmonic wavelet packet transform. *Journal of Southeast University: Natural Science Edition* 38 (6): 996–999.
3. Wu, Guo feng, and Zhang Wen yuan. 2000. The brain waves generated by neural physiological mechanism. *Journal of Clinical Electroencephalogram* (3): 188–190.
4. Haufe, S, Treder, M. S., Gugler, M. F. et al. 2011. Frontiers|EEG potentials predict upcoming emergency braking during simulated driving. *Journal of Neural Engineering* 8(5).
5. Su, Jing. 2006. The mental disease diagnosis system based on brain wave biological database. Hua Zhong University of Science and Technology.
6. Yin, Yan hong. 2008. Experimental study of driver fatigue simulation based on brain waves and blinking. Tongji University.
7. Liu, Qing jie, Chen, Gui ming, Liu Xiaofang. et al. 2009. FFT and wavelet in signal denoising based on data acquisition and processing. 24 (z1): 58–60.

8. Wang, Hao wen, Qian, Zhi yu, Li, Hong jing. et al. 2013. Analysis of the effect of EEG and wavelet transform on the information separation of FFT signal beta wave. *Journal of Biomedical Engineering* (4): 704–709.

9. Liu, Yang, Zhengbing, He, Wei Guan. et al. 2016. Exploring the relationship between EEG and ordinary driving behavior.

10. Ji, Zhong, Qin, Shu ren. 2003. EEG feature extraction of non-stationary signal analysis technology. Proc of China Aviation Institute of signal and information processing branch of the second joint academic exchange meeting.

11. Wu, Xiao bin, and J. Qiu 2004. Research progress of EEG signal analysis and processing method based on time frequency analysis. *International Journal of Biomedical Engineering* 27 (6): 321–326.

12. Zhang, Chong, Zheng, Chong xun, Ouyang Yi. et al. 2008. Mental fatigue analysis based on the characteristics of EEG power spectrum. *Aerospace Medicine and Medical Engineering* 21 (1): 35–39.

13. Chen, A. C. N., Wei jia. F., Hui xuan, Z. et al. 2008. EEG default mode network in the human brain: Spectral regional field powers. *Neuroimage* 41(2):561–74.

14. Jun, Feng Gao, F, and Wang Pei. 2010. Based on independent component analysis (ICA) and manifold learning EOG artifact removal. *Journal of Xi'an Jiaotong University* 44 (2): 113–118.

15. Wu, Xiao pei, Feng, Huan qing, Zhou, He qin. et al. 2004. The EEG feature extraction based on wavelet transform and independent component analysis. *Journal of Instrumentation* 25 (1): 116–120.

Evaluating the Effectiveness of Speed Bumps: An Empirical Study in Campus

Jing Lv, Cheng Li, Hongwei Guo, Wuhong Wang and Xiaobei Jiang

Abstract The installation of speed bumps on the roads aims to improve traffic safety by forcing vehicle to decelerate. The actual deceleration effectiveness of speed bump is closely related to driving behaviors, installation place, dimension, structure of speed bump, etc. Through video capture in campus and urban roads and coordinate transformation, the spot velocity curves of each vehicle's traversing the speed bump are obtained. Based on the trajectory data, evaluations on the actual performance of speed bump are carried out. Results show that speed bumps can have certain effect on speed reduction. But the disperse distribution of spot velocities indicates diverse driving behaviors, meaning that speed bumps may not influence all of the vehicles.

Keywords Speed bump · Deceleration effectiveness · Video capture · Driving behaviors' diversity

Introduction

Speed bump is considered to be one of the typical traffic calming measures and can provide favorable effectiveness in speed control. However, due to some impractical arrangements, the speed control results of deceleration bumps in Chinese cities are sometimes far from satisfactory. Hence, a scientific analysis is needed to evaluate the speed calming effectiveness of deceleration bumps, and the optimization

J. Lv · C. Li · H. Guo (✉) · W. Wang (✉) · X. Jiang
School of Mechanical Engineering, Beijing Institute of Technology,
Beijing 100081, China
e-mail: guohongwei@bit.edu.cn

W. Wang
e-mail: wangwuhong@bit.edu.cn

J. Lv · C. Li · H. Guo · W. Wang · X. Jiang
School of Management and Economics, Beijing Institute of Technology,
Beijing 100081, China

designs can resultantly proceed [1]. When applied in urban roads or residential areas, speed bumps produce positive effects along with negative effects. Vehicular speeds and traffic accidents are reduced, whereas possible noise pollution or exhaust pollution may be increased. Many scholars not only study the speed reduction effectiveness of speed bumps, but also take traffic and environmental performance into considerations. In comparison with other traffic calming measures, comprehensive and detailed evaluations on speed bumps are obtained. Pau [2] studied 23 speed bumps installed in Italy and found that the employment of such a device is not really effective in protecting crosswalks. One third of the cases the 85th‰ of speed measured at the speed bumps is higher than the posted speed limit (50 km/h). Mountain [3] conducted a comparison between speed management schemes in Britain, including safety cameras, horizontal schemes (such as roundabouts and chicanes), and vertical schemes (such as speed bumps and speed cushions). Results showed that engineering schemes incorporating vertical deflections offered the largest and most consistent safety benefits. Through a three-month study on speed bumps in the city of Omaha, Gorman [4] found that the devices influence the 85% speed in a statistically significant, but do not indicate a significant increase or decrease in accident statistics. Lee [5] compared the effectiveness of various traffic calming measures using a framework incorporating traffic performance and safety, and environmental and public health impacts. The results showed that speed bumps were in a medium level but had the best performance in vehicle emissions. Besides speed control effectiveness and comparison with other traffic calming measures, researchers also performed evaluations on speed bumps in the perspective of vehicle vibration, traffic simulations, etc.

However, most of the literatures did not involve a process analysis of vehicles independently passing through the speed bump and possible distractions brought by other traffic participants. Additionally, the speed calming effectiveness is closely related to the installation sites and manners, so it is necessary to assess the influence comprehensively in terms of elements and settings. Resultantly, we are able to provide grounded basis for improvements of traffic calming measures.

The rest of this paper is organized as follows: Firstly, the data collections are specified; then variables concerning the process of vehicles' traversing speed bumps are analyzed; and finally, evaluations on effectiveness of speed bumps are presented and conclusions are given in the last section.

Data Collection

The installation of speed bumps should satisfy the necessary conditions and certain standards, and take traffic impact into account. So that it is advised to collect traffic states data such as traffic volume and traffic load, and road dimension data such as lane width, parking lane width, and speed limit. This paper mainly focuses on speed reduction effectiveness; thus, the data collected from field investigations are sequential spot speeds during the process when vehicles approach, traverse,

and departure from the speed bump. Besides, road dimension data and traffic volumes will be measured.

Common installation places of speed bumps are entrance, residential and school sections. Since we aim to collect the speed data when vehicles drive straight through the speed bumps, entrances and ending branch of T-shape intersections are excluded in the list of observation sites. We choose four investigation sites including roads in campus and urban areas where speed bumps are installed.

In order to obtain the velocity curve of the vehicles, it is required to output a series of continuously varying velocity values, or a discrete value composed of smaller intervals. Hence, video capture methods are comparatively superior to conventional methods such as speed radar, manual measurements, or auto counters. The video capture method is illustrated in Fig. 1. Coordinate transformations are adopted to obtain the real coordinates of vehicles (see Eqs. 1, 2).

Coordinate transformation formulas are listed as follows: Four reference points with known screen coordinates and real coordinates are required to calibrate eight parameters $(C_1 - C_8)$ in the formulas.

$$X_r = \frac{C_1 + C_2 X_s + C_3 Y_s}{C_4 X_s + C_5 Y_s + 1} \tag{1}$$

$$Y_r = \frac{C_6 + C_7 X_s + C_8 Y_s}{C_4 X_s + C_5 Y_s + 1} \tag{2}$$

where $X_r(Y_r)$ is the $X(Y)$-coordinate in real world, $X_s(Y_s)$ is the $X(Y)$-coordinate on screen.

With certain post-processing techniques on coordinated data extracted from raw video, we are able to obtain spot speeds that are continuously varying. Speed of

Fig. 1 Schematic illustration of video capture method
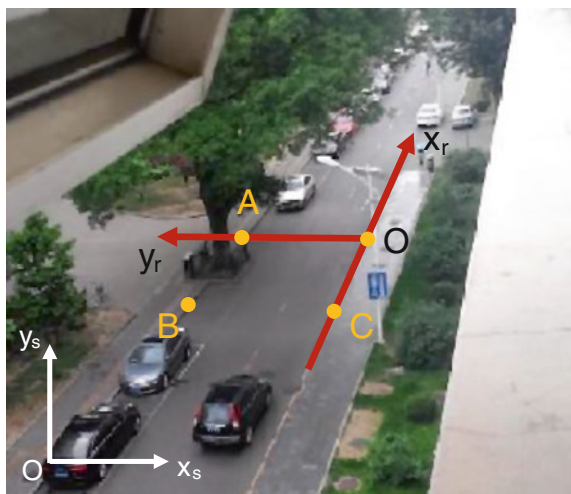
vehicles is calculated by the ratio of displacement and time intervals, as presented in Eq. 3.

$$v_t = \dot{x}_t = \frac{\vec{x}_{t+\Delta t} - \vec{x}_t}{\Delta t} = \frac{\vec{x}_{f+\Delta f} - \vec{x}_f}{\Delta f / fr} \qquad (3)$$

where v_t is the velocity at time t; $\vec{x}_t$ is the position coordinates at time t; $\vec{x}_f$ is the position coordinates at time t represented by frame index; fr is the frame rate of the video.

As shown in Eq. 3, time intervals are calculated as the ratio of frame intervals and frame rate. Videos recorded by our camera have a frame rate of 30 fps, so that we have already addressed the acquisition method of time stamps and time intervals. Afterward, driving behaviors of vehicles will be analyzed based on the speed data.

Data Analysis

This paper mainly focuses on the driving behavior of the vehicle passing the speed bump in an independent state; thus, interference from pedestrians or non-motor vehicles should be as little as possible. Moreover, data filtering should fully consider data availability and reliability, resulting in a total number of 79 samples at four sites. The number of sample at each site is shown in Table 1.

In the analysis of spot speed, this paper does not select the spot speed in the general sense, which is the coordinate points within a certain distance range of the installation point of the speed bump. In contrast, we incorporate driving behaviors into the choice of spot speeds. More specifically, we use the distance between vehicle and speed bump as reference points and capture the deceleration process of vehicles in both directions. The statistical diagrams of spot speed are shown in Fig. 2.

It is shown in Fig. 1 that the slowdown effect of the speed bump is indicated by the average velocity change at the observation site. Comparing the average initial velocity and the average velocity when vehicles traverse the speed bump, we find that, at Site 1, the velocity is reduced from 9.13 m/s (32.87 km/h) to 6.27 m/s (22.57 km/h), indicating a decrease by 31.3%. Additionally, the average post-bump velocity returns to 84.3% of the average initial velocity. As for sites 2, 3, 4, the decreasing amplitude of velocity is approximately within the range of 20–25%. Respectively, velocity at the speed bump decreases by 19.63, 25.16, and 22.95% for sites 2, 3, 4.

Table 1 Number of samples at each investigation site

Site ID	1	2	3	4
Number	24	21	20	18
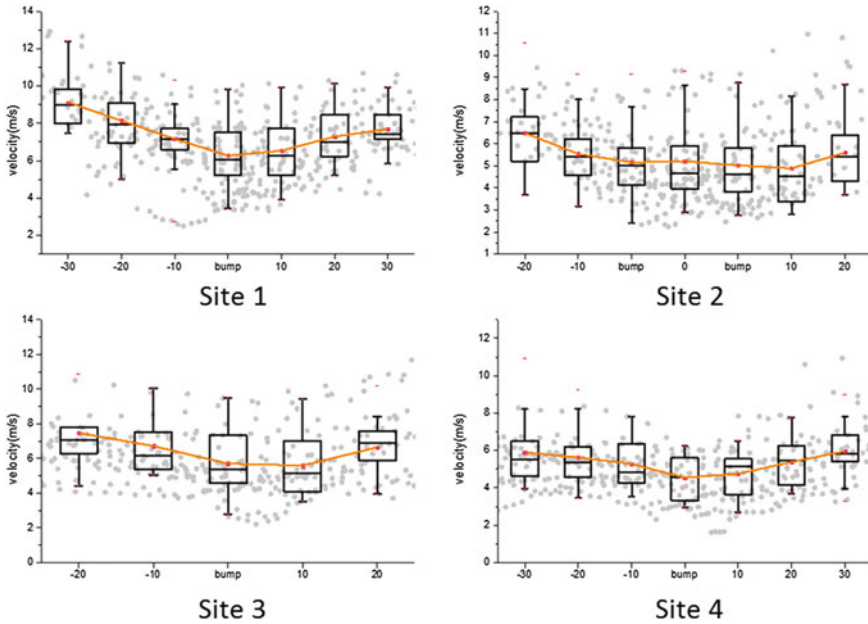
Fig. 2 Statistical diagrams of spot speed during deceleration process

Table 2 Spot speed variance at each investigation site (unit: m^2/s^2)

Location	−30	−20	−10	Bump	10	20	30
Site ID: 1	1.882	2.253	2.096	3.065	2.886	2.131	1.202
Site ID: 2	2.711	2.158	2.772	3.204	2.970	2.984	2.331
Site ID: 3	–	3.584	2.448	3.292	2.852	2.775	–
Site ID: 4	3.471	2.394	1.905	1.220	1.629	1.478	2.213

The variance of spot speeds at each observation site is listed in Table 2. It can be seen that the variance of spot speed shows a trend of increasing first before decreasing, indicating difference driving behaviors between vehicles when passing through speed bumps. The existence of this phenomenon reflects that some of the vehicles may not slow down when traversing speed bump, which is namely deceleration rate. Generally, greater deceleration rate represents that more vehicles are influenced by speed bumps, and, as a result, speed bump is more effective in speed control.

Among 79 samples, direct outputs are initial velocity (IV), deceleration range (DR), and passing speed (PS). Initial velocity is the speed where deceleration begins. To specify, when the curve of spot speed shows significant decline, the very distance from speed bump is the deceleration range and the very spot speed is the

initial velocity. Other related variables are deceleration amplitude [absolute (ADA) and relative (RDA)], average deceleration (AD), deceleration time (DT), and ratio of passing speed to initial speed (PS/IV). Calculation formulas of related variables are listed as follows:

$$\text{Absolute deceleration amplitude (ADA)} \quad \Delta v^* = v_0 - v_p \qquad (4)$$

$$\text{Relative deceleration amplitude (RDA)} \quad \delta = \frac{\Delta v^*}{v_0} \times 100\% \qquad (5)$$

$$\text{Average deceleration (AD)} \quad \bar{a} = \frac{v_0^2 - v_p^2}{2S_{\text{dec}}} \qquad (6)$$

$$\text{Deceleration Time (DT)} \quad T_{\text{dec}} = \frac{\Delta v^*}{a} \qquad (7)$$

where v_0 is the initial velocity (IV); v_p is the passing speed (PS); S_{dec} is the deceleration range (DR).

Table 3 shows the analysis of the correlation coefficients of the eight variables. Through the preliminary correlation analysis, we can see that there is a large positive correlation between the velocity and the initial velocity, and the absolute deceleration amplitude also has a positive correlation with the initial velocity. But the ratio of passing speed to initial velocity has extremely weak correlation between the relative deceleration amplitude and the initial travel speed.

Deceleration Effectiveness Evaluation

Evaluation indicators of the effectiveness of speed bump are mainly speed (including spot speed and average speed), speed variance, deceleration, and deceleration variance. Speed represents overall deceleration effect and general safety trend, while spot deceleration reflects micro-safety, and speed variance and deceleration

Table 3 Correlation coefficients of the eight variables

	IV	DR	PS	PS/IV	ADA	RDA	AD	DT
IV	1							
DR	0.424	1						
PS	0.707	0.063	1					
PS/IV	−0.016	−0.357	0.676	1				
ADA	0.523	0.507	−0.233	−0.837	1			
RDA	0.016	0.357	−0.676	−1	0.837	1		
AD	0.652	−0.055	0.162	−0.404	0.702	0.403	1	
DT	−0.171	0.780	−0.465	−0.524	0.326	0.524	−0.363	1

variance show driving behaviors' diversity. In fact, the ideal deceleration effect of speed bump is that spot speed declines significantly as the vehicle approaches the speed bump. After passing through the speed bump, vehicle shall maintain low velocity for another period of time. From the macroscopic point of view, the speed and acceleration do not show too large fluctuations, meaning that drivers tend to take some deceleration action when passing through the speed bump.

We have obtained favorable results of deceleration effectiveness of speed bump from empirical data. The passing speed of vehicles is comparatively low at some sites, and the low velocity maintains for some time. In the meanwhile, speed variance shows preferable trend, namely a 'V' shape among the observation spots. Although the average speed can experience large decline, spot speed variance shows an inverse trend. It is explained as that driving behaviors' diversity leads to disperse distribution of passing speed. There is a certain percentage of the driver ignoring the speed bump, making the actual impact of the slowdown ridge effect greatly reduced. Besides, accident risk in the downstream of speed bump has not been effectively controlled.

Conclusions

Based on the velocity curves obtained through video acquisition and coordinate transformation, this paper conducts a quantitative analysis on the velocity variation trend, aiming to evaluate actual effectiveness of speed bumps. Results show that average spot velocity decreases significantly at the speed bump, but the difference between vehicles is obvious and speed bumps may not have impact on all vehicles. The actual effect of speed bump is closely related to the setting mode, the driving behaviors, and the structure. When employing speed bumps, it is advised to follow the deceleration characteristic of the device so that traffic safety can be effectively improved. Due to the limitation of sample size, the analysis of deceleration effect does not involve comparisons between different locations and different sizes. In future work, the number of samples will be further increased and the effectiveness of speed bumps will be studied systematically.

References

1. Ewing, Reid. 1999. Traffic calming: State of the practice. *Regulatory Constraints*.
2. Pau, M., and S. Angius. 2001. Do speed bumps really decrease traffic speed? An Italian experience. *Accident Analysis and Prevention* 33 (5): 585–597.

3. Mountain, L.J., W.M. Hirst, and M.J. Maher. 2005. Are speed enforcement cameras more effective than other speed management measures? The impact of speed management schemes on 30 mph roads. *Accident Analysis and Prevention* 37 (4): 742–754.
4. Gorman, Michael N., M. Moussavi, and P. T. Mccoy. 1989. Evaluation of speed hump program in the city of Omaha. *ITE Journal 59*.
5. Lee, Gunwoo, et al. 2013. An evaluation framework for traffic calming measures in residential areas. *Transportation Research Part D Transport & Environment* 25. 8: 68–76.
6. Antić, Boris, et al. 2013. The influence of speed bumps heights to the decrease of the vehicle speed—Belgrade experience. *Safety Science* 57. 57: 303–312.
7. Kokowski, Piotr, and R. Makarewicz. 2006. Predicted effects of a speed bump on light vehicle noise. *Applied Acoustics* 67 (6): 570–579.
8. Meng, Jianping, and Jiefang Zhang. 2006. The effect of deceleration strips upon traffic flow. *Modern Physics Letters B* 20 (14): 835–841.

Traffic Signal Optimization Based on System Equilibrium and Bi-level Multi-objective Programming Model

Xiao-ting Wang, Yu-lin Chang and Peng Zhang

Abstract Based on the design principle of minimizing the automotive exhaust emissions and the total impedance of the road network, an urban road traffic signal control method of bi-level multi-objective programming model is established by designing the heuristic particle swarm optimization (PSO) algorithm. First of all, the upper-level model which combined the vehicle emissions model and system optimum assignment model is built, and the lower-level model is built based on minimizing the sum of the link travel time function integral. Then, the heuristic PSO algorithm is designed and transformed to solve upper-level and lower-levels model iteratively by two PSO algorithms. Ultimately, by altering the weight parameters of the upper model, the model is dealt with separately in case of single target and multi-target, the optimization results of which is compared with the VISSIM simulation results and the optimization results by means of heuristic genetic algorithm. The simulation results show that bi-level multi-objective control method, which could improve the operating quality of road network, is of great optimization ability and can effectively reduce the automotive exhaust emissions and the total impedance of the road network.

Keywords Traffic engineering · Bi-level multi-objective programming model · PSO algorithm · Signal control · System equilibrium

Introduction

Air pollution caused by the fast development of traffic motorization has become one of the main factors affecting the urban environmental quality. Bi-level programming model has been an effective and significant modeling tool in the field of

X. Wang (✉) · Y. Chang · P. Zhang
School of Automotive and Traffic Engineering, Jiangsu University, Zhenjiang 212013, China
e-mail: wxt412c@163.com

Y. Chang
Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing 211189, China

traffic planning and traffic management, and it has a considerable degree of complexity due to containing the master–slave hierarchical decision-making process, which causes the upper-level and lower-level models to have their own decision variables, constraint conditions, and objective functions. Therefore, the optimization results of the upper-level model not only depend on the upper-level model decision variables, but are also affected by the optimal solutions of the lower-level model that are closely related to the upper decision variables. This decision-making model makes the entire system to get the best benefits during the process of cooperation and coordination, but can also make decisions freely with mutual constraints. Based on the design principle of stochastic user equilibrium, Michael J. Maher forecasted the path choice of congested road network and optimized the signal timing program of intersection by building bi-level optimizing model [1]. Sun [2] established a bi-level programming model on stochastic route choice and time-dependent demand in road network and developed heuristic solution approach consisting of genetic algorithm and cell transmission simulation-based incremental logit assignment procedure. Based on the route choice of users, Zhi-lin Lv developed a bi-level programming model to minimize the pollution emissions in urban expressway network and maximize traffic volumes entering the network [3]. Zhou [4] studied traffic signal cycle length and green ratio by taking environment factors into consideration, proposing a bi-level programming model based on user equilibrium route choosing model. Gao [5] optimized intersection signal control parameters, departure frequency, and routes of evacuation vehicle at assembly points through establishing bi-level programming model. In the terms of road coordinate control, Jian-min Xu built a urban arterial road coordinate control bi-level programming model by introducing group decision-making theory into conventional traffic signal control method, which could reduce delay and improve segment travel speed effectively [6]. Based on the regional coordinate signal control theory and user equilibrium, Yue Hu put forward a bi-level programming model for traffic assignment and signal control optimization [7]. Currently, the existing literature about bi-level programming model is just to optimize delay, stops, and environmental benefits in road networks, not to verify whether the total impedance of network to achieve optimal results under user optimal distribution.

Therefore, aiming at user route choosing problem under dynamic network signal control, this paper proposes bi-level multi-objective programming model consisting of traffic control and user route choice. The upper-level model is to minimize the system total impedance based on optimizing vehicle exhaust emissions, and the lower-level model for user route choice is established based on user equilibrium. Eventually, the bi-level multi-objective programming model is put forward on system equilibrium path selection model. In addition, this paper solves the model by designing heuristic particle swarm algorithm, which is compared with heuristic genetic algorithm, and is of great superiority and optimization ability.

Bi-level Programming Problem

Bi-level programming problem can be described as impartial combination games between traffic planners and road users. Therefore, the mathematical forms of bi-level programming problem are as follows:

$$
\begin{cases}
\min_{x \in X} P(x, y) \\
\text{s.t. } G(x, y) \leq 0 \\
\min_{y \in Y} p(x, y) \\
\text{s.t. } g(x, y) \leq 0
\end{cases}, \tag{1}
$$

where $P(x, y)$ and $G(x, y)$ represent separately the objective function and constraints of the upper-level model, in which $x \in X \subset R^n$ and $y \in Y \subset R^m$ are the decision variables of upper-level and lower-level models, and m and n are integers. Likewise, $p(x, y)$ and $g(x, y)$, respectively, represent the objective function and constraints of the lower-level model, $G(x, y) \leq 0$ and $g(x, y) \leq 0$ are the equality and inequality constraints for variables, respectively. On the whole, this mathematical model takes the variable x of the upper-level issue as control parameters, optimizing the objective functions of the lower-level model.

Bi-level Multi-objective Programming Model

The Upper-Level Model

Since bi-level multi-objective programming model can more thoroughly convey the wishes of managers, the objective function of the upper-level model is to minimize the vehicle exhaust emissions in network and the total impedance of road users.

Firstly, the vehicle exhaust emissions in network are the summations of the exhaust emissions when the vehicle travels on the road and idling emissions at the intersection. As a result, the objective functions to make the least motor vehicle exhaust emissions in network are:

$$
\min : E_s = \alpha_1 E_s' + \alpha_2 E_s'', \tag{2}
$$

$$
E_s' = \sum_{i \in I} EF_{i,d}^{PCU} \cdot x_i \cdot d_i, \tag{3}
$$

$$
E_s'' = \sum_{i \in I} EF_i^{PCU} \cdot x_i \cdot L_i. \tag{4}
$$

where E_S is the total automotive exhaust emissions in the network, E_s' is the exhaust emissions when the vehicle travels on road, and E_s'' is idle speed emissions at

intersection. α_1 and α_2 is, respectively, the weight of the exhaust emissions at the intersection and road segments; $EF_{i,d}^{PCU}$ and EF_i^{PCU}, respectively, are the exhaust emissions factors under idle and driving conditions. d_i is the travel delay on road segment i, which is given by $d_i = \frac{C(1-\lambda)^2}{2\left(1-\frac{x_i}{S}\right)} + \frac{\left(\frac{x_i}{\lambda S}\right)^2}{2x_i\left(1-\frac{x_i}{\lambda S}\right)}$, in which C is the signal cycle, λ the splits, and S the saturating flow. x_i is the traffic on road segment i, and L_i is the length of road segment i.

Subsequently, road manager is always willing to minimize the total impedance that is the total travel time, making the system to achieve equilibrium; so the objective function equals to system optimal allocation model and is given by:

$$\min : \ Z'(X) = \sum_i x_i t_i(x_i)$$
$$\text{s.t.} \begin{cases} \sum_k f_k^{rs} = q_{rs} \\ f_k^{rs} \geq 0 \quad \forall r, s, i \end{cases} , \tag{5}$$

where $t_i(x_i)$ is the impedance function of route, f_k^{rs} is the traffic of route k between the junctions r and s, q_{rs} is the OD volume between the junctions r and s. Due to the intersection delay, the impedance function of route is divided into two parts, namely the travel time on road and intersection delay, which is:

$$t_i(x_i) = t_i'(x_i) + D_i, \tag{6}$$

from which we use BPR function developed by US Bureau of Road to define the travel time on road, and its mathematical form is:

$$t_i'(x_i) = t_i(0)\left[1 + \alpha(x_i/S)^\beta\right], \tag{7}$$

where $t_i(0)$ is the travel time required when there is none vehicle on the road; α and β are the parameters to be calibrated of function and are supposed to be 0.15 and 4.

To sum up, the upper-level multi-objective model is as follows:

$$\min : \ F(X) = \alpha\left(\alpha_1 \sum_{i \in I} EF_{i,d}^{PCU} \cdot x_i \cdot d_i + \alpha_2 \sum_{i \in I} EF_i^{PCU} \cdot x_i \cdot L_i\right) + \beta \sum_i x_i t_i(x_i)$$
$$\text{s.t.} \begin{cases} \sum_k f_k^{rs} = q_{rs} \\ f_k^{rs} \geq 0 \quad \forall r, s, i \\ C \geq 30, \lambda \geq 0.05 \\ C \leq 120, \lambda \leq 0.95 \end{cases} ,$$
$$\tag{8}$$

where α and β are weighting coefficients; during the traffic planning process, the value of which denotes, respectively, the impact of two goals, namely minimum emissions and the total independence, and is given by $\alpha + \beta = 1$.

The Lower-Level Model

During the model establishing process, the total impedance of road users in network is not only taken into thought, but the path choice is also taken into consideration to make user behavior to meet the requirements of system and user equilibrium simultaneously. Therefore, the lower-level model, namely traffic assignment model, is given by:

$$
\begin{cases}
\min: \ Z(X) = \sum_i \int_0^{x_i} t_i(x_i)\mathrm{d}x_i \\
\text{s.t.} \begin{cases} \sum_k f_k^{rs} = q_{rs} \\ f_k^{rs} \geq 0 \quad \forall r, s, i \end{cases}
\end{cases} \tag{9}
$$

The objective function of the lower-level model by taking Eqs. 6 and 7 into Eq. 9 and integrating is as follows:

$$
Z(X) = \sum_i \left[t_i(0)x_i + 0.03S \cdot t_i(0)\left(\frac{x_i}{S}\right)^5 + 0.5SC(1-\lambda)^2 \cdot \ln\left(1 - \frac{x_i}{S}\right) - \frac{x_i}{2\lambda S} - \frac{1}{2}\ln(x_i - \lambda S) \right].
$$
$$\tag{10}$$

Bi-level Multi-objective Programming Model

In summary, the bi-level multi-objective programming model is built to show:

$$
\begin{aligned}
U: \ & \min: \ \alpha\left(\alpha_1 E_s' + \alpha_2 E''\right) + \beta \sum_i x_i t_i(x_i) \\
& \text{s.t.} \begin{cases} \sum_k f_k^{rs} = q_{rs} \\ f_k^{rs} \geq 0 \quad \forall r, s, i \\ C \geq 30, \lambda \geq 0.3 \\ C \leq 120, \lambda \leq 1 \end{cases} \\
D: \ & \min: Z(X) = \sum_i \int_0^{x_i} t_i(x_i)\mathrm{d}x_i \\
& \text{s.t.} \begin{cases} \sum_k f_k^{rs} = q_{rs} \\ f_k^{rs} \geq 0 \quad \forall r, s, i \end{cases}
\end{aligned} \tag{11}
$$

The objective function of the upper-level problem on the above is to minimize the automotive exhaust emissions in network and the total impedance; for each road user, the lower-level model of which is to minimize the travel cost.

Heuristic Particle Swarm Optimization Algorithm

In 1995, Kennedy and Eberhart were inspired by birds foraging behavior to propose particle swarm optimization algorithm, namely PSO algorithm, which starts from a random solution and finds the optimal solution of problem by iterating and is of great ability to search the optimal solution in a global variable based on the current solution [8]. Zhao [9] solved bi-level programming problem based on layer iteration by advanced PSO algorithm, calculating the upper-level problem by means of PSO algorithm and the lower-level problem with conventional methods, then to obtain the global optimal solutions by layer iteration. Based on two metamorphic PSO algorithm, Xiang-yong Li dealt with the upper-level and lower-level problems by interactive iteration, then to solve common bi-level programming problem directly by simulating the decision making process of bi-level programming model [10]. Thus, we calculate, respectively, the upper-level and lower-level problems via using PSO algorithm and acquire the best solution by means of interactive iteration. The bi-level algorithm is outlined as follows:

Step 1: Argument initialization. Initialize the size of particle swarm N, inertia weight w, learning factor c_1, c_2, maximum iteration M, the dimensions of bi-level problem D, and set $k = 1$.

Step 2: According to the constraints of the upper-level problem, the initial solution x_0 of upper-level decision variable x can be obtained and set $fbest = \text{INF}$.

Step 3: According to the initial solution x_0, the corresponding optimal solution y_0 of the lower-level problem can be calculated using PSO algorithm.

Step 4: Return the optimal solution y_0 to the upper-level problem and get the optimal solution and individual fitness with PSO algorithm, then to get the solution x and target function value f. Otherwise, for the constraints, the fitness function is formed by means of penalty function.

Step 5: Compare the target function value f with $fbest$, if $f < fbest$, then set $fbest = f$, $xbest = x$, $ybest = y_0$, or else go to Step 6.

Step 6: If the termination condition is met, stop; otherwise, set $x_0 = xbest$, $k = k + 1$ and go to Step 3.

Step 7: Output the optimal solution x, y, and target function value $xbest$, $ybest$ of the upper-level and lower-level modes.

Numerical Application

The proposed heuristic PSO algorithm is implemented in a simple network as shown in Fig. 1. There are two O–D pairs: from O1 to D1 and from O2 to D2. There are seven alternative paths for O–D pairs and each O–D traffic flow is 1000 veh/h. In this simple network, there are two signalized intersections A and B, and a straight

Fig. 1 Structure diagram of road network
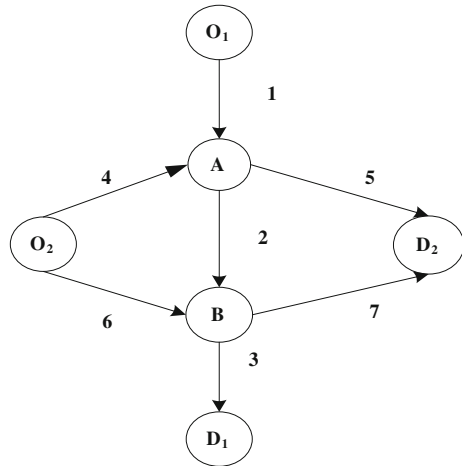
forward control plan for each signalized intersections is maken in case of one-way traffic. T_A and T_B ($T_i \in [30, 120]$), respectively, represents the cycle length of the signalized intersections A and B. We suppose that $l_1 = l_3 = 500$ m, $l_2 = 300$ m, $l_4 = l_7 = 400$ m, and $l_5 = l_6 = 600$ m. According to Fig. 1, we assume that the relationship of green ratio at signalized intersection A is: $\lambda_1 + \lambda_4 = 1$, $\lambda_i \in [0.05, 0.95]$. Set the size of particle swarm $N = 200$, inertia weight $w = 0.6$, learning factor $c_1 = c_2 = 1.5$, maximum iteration $M = 200$, the dimensions of bi-level problem $D = 7$.

Because the mathematical form of different pollutants is similar, based on different automotive exhaust emissions, the assignment is alike by balancing and distributing traffic [11]. Otherwise, due to the high proportion of CO in all exhaust emissions, this numerical example only analyzes the effects of CO pollutants on traffic assignment.

$\alpha = 0, \beta = 1$

To rule out the single correlation between "the minimum total impedance of all road users" and "user optimal assignment," set the parameters α and β on the upper-level problem: When α is zero and β is one, the upper-level problem can be simplified into single-objective programming problem. Meanwhile, set initialization: $T_A = T_B = 80$, $\lambda_1 = \lambda_4 = \lambda_2 = \lambda_6 = 0.5$.

Based on MATLAB software, write project on heuristic PSO algorithm, and the callable function is written by handling the target function and constraints of the upper-level and lower-level models with penalty function. Using the above heuristic PSO algorithm, after repeated checking, this paper takes assignment on segment 1, for example, to draw the flow curve of the upper-level and lower-level models, whose running results are shown as follows:

Fig. 2 Comparison of distribution flow in the road segment 1
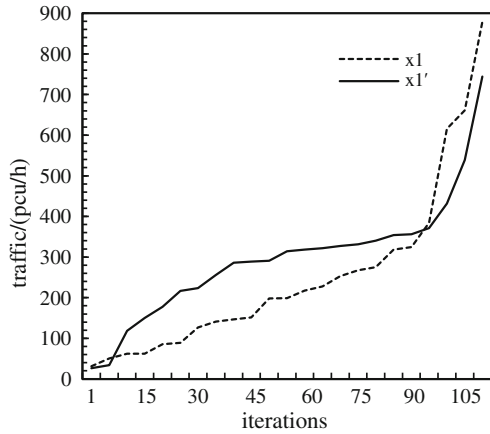
$x_1 = 234.255$, $x_2 = 661.303$, $x_3 = 371.287$, $x_4 = 126.982$, $x_5 = 327.382$, $x_6 = 227.538$, $x_7 = 177.345$. $T_A = 80.228$, $T_B = 80.831$, $\lambda_1 = 0.666$, $\lambda_4 = 0.334$, $\lambda_2 = 0.805$, $\lambda_6 = 0.195$, $Z'(X) = 535,080$ (Fig. 2).

Assuming that the single correlation between "the minimum total impedance of all road users" and "user optimal assignment" exists, after each iteration, the assignment results for road segment 1 should be a trend in the same and complete coincidence curve. However, it can be obtained from Fig. 2, in the process of solving this bi-level programming model, which is the upper-level and lower-level model acquire their optimal solutions by interactive iteration, then to obtain the optimal solutions of this model.

$\alpha = 0.5, \beta = 0.5$

In order to make the upper-level model to become a multi-objective optimization model, set parameters α and β, when α is 0.5 and β is 0.5. Meanwhile, set initialization: $T_A = T_B = 80$, $\lambda_1 = \lambda_4 = \lambda_2 = \lambda_6 = 0.5$. Based on MATLAB software, using the above heuristic PSO algorithm, after repeated checking, the running results are shown as follows:

$x_1 = 234.255$, $x_2 = 561.303$, $x_3 = 371.287$, $x_4 = 226.982$, $x_5 = 257.382$, $x_6 = 207.538$, $x_7 = 247.345$. $T_A = 90.228$, $T_B = 85.831$, $\lambda_1 = 0.566$, $\lambda_4 = 0.434$, $\lambda_2 = 0.705$, $\lambda_6 = 0.295$, $F(X) = 455,080$, $E(X) = 405,080$, and $Z'(X) = 485,080$.

The Comparison of Optimization Results

This model is calculated, respectively, by heuristic PSO algorithm and heuristic genetic algorithm, which is also simulated by VISSIM to obtain the exhaust emissions and the total impedance, and the comparisons of them are shown in Table 1.

Table 1 Comparison of the model algorithm optimal results

	α, β	E/g	$Z'(X)$/g	Optimization rate of emissions (%)	Optimization rate of the total independence (%)
VISSIM simulation results		689,020	602,340		
Heuristic PSO	$\alpha = 0,$ $\beta = 1$		535,080		11.17
	$\alpha = 0.5,$ $\beta = 0.5$	405,080	485,080	41.21	19.47
Heuristic genetic algorithm	$\alpha = 0,$ $\beta = 1$		566,230		5.99
	$\alpha = 0.5,$ $\beta = 0.5$	420,380	510,250	38.99	15.29

The above table shows that the signal cycle length with heuristic PSO algorithm is 90 and 85 s, and green ratio is 0.57, 0.43, 0.70, and 0.30, which reduces the total automotive exhaust emissions in the network by 41.21% and the total impedance by 16.15% by comparing with the VISSIM simulation results. Compared with heuristic genetic algorithm, both based on the optimization results of the single-objective model or multi-objective model, the optimization results using heuristic PSO algorithm have obvious superiority, which increase, respectively, the optimization rate of the exhaust emissions and the total impedance by 2.22 and 4.18%.

Summary

The problem of combined trip matrix estimation and system equilibrium assignment has been addressed in this paper. The proposed bi-level multi-objective programming model based on system equilibrium describes the automotive exhaust emissions problem in urban network and road manager's estimation of travel time as well as user's, in which the upper-level model minimizes the exhaust emissions and the total impedance from the perspective of manager; on the other hand, from road user perspective, the lower-level model establishes user equilibrium assignment model. At the same time, this paper designs heuristic PSO algorithm to solve this model, and the numerical example shows the great feasibility of the model and superiority of heuristic PSO algorithm compared with heuristic genetic algorithm.

References

1. Michael, J.M., X.Y. Zhang, and V.V. Dirck. 2001. A bi-level programming approach for trip matrix estimation and traffic control problems with stochastic user equilibrium link flows. *Transportation Research Part B* 35: 23–40.

2. Sun, D.Z., F.B. Rahim, and S.T. Waller. 2006. Bi-level programming formulation and heuristic solution approach for dynamic traffic signal optimization. *Computer-Aided Civil and Infrastructure Engineering* 21: 321–333.
3. Lv, Z.L., B.Q. Fan, and J.J. Liu. 2006. Bi-level multi-objective programming model for the ramp control and pollution control on urban expressway networks. *Control and Decision* 21 (1): 64–67, 87.
4. Zhou, S.P., C.Z. Wu, and X.P. Yan. 2009. Bi-level optimization model for signal timing of intersections under environmental objective. *Journal of Wuhan University of Technology (Transportation Science & Engineering)* 33 (4): 715–717, 721.
5. Gao, M.X. 2014. Optimization of vehicle departure arrangement and traffic control in traffic evacuation based on bi-level programming. *Chinese Journal of Management Science* 22 (12): 65–71.
6. Xu, J.M., and S.Y. Chen. 2010. Traffic signal control optimization based on group decision making theory and bi-level programming model. *Journal of Highway and Transportation Research and Development* 27 (12): 134–139, 158.
7. Hu, Y. 2015. *Research on traffic flow assignment and signal control technology for urban road*. Hangzhou: Zhejiang University.
8. Kennedy, J., and R.C. Eberhart. 1995. Particle swarm optimization. In *Proceeding of IEEE International Conference on Neural Network*. IEEE Service Center, Piscataway, NJ, 1942–1948.
9. Zhao, Z.G, X.Y. Gu, and T.S. Li. 2007. Particle swarm optimization for bi-level programming problem. *System Engineering-Theory & Practice* 8: 92–98.
10. Li, X.Y., P. Tian, and X.P. Min. 2006. A hierarchical particle swarm optimization for solving bi-level programming problems. Lecture notes in computer science, Zakopane, Poland, 25–29 June 2006, 4029, pp. 1169–1178.
11. Zhao, T., and Z.Y. Gao. 2003. The bi-level programming models in the urban transport network design problem. *China Civil Engineering Journal* 36 (1): 6–10.

Research on Fundamental Solutions to Curb Parking Problems in City

Di Wang, Xue Han and Guixian Xing

Abstract It is a systematic project to fundamentally solve the curb parking problem in city. It works temporarily to increase parking management enforcement, perfect road parking sign or strengthen travelers' thought education. From perspective of demand management for curb parking, this paper puts forwards five fundamental solutions including making perfect the city transportation system, improving travelling structure, and taking advantage of technological intelligence to curb parking problems.

Keywords Solutions · Curb parking · Parking and ride · Parking APP

Background

Compared with appertaining parking and public parking lot, influence of curb parking on transportation is most serious. Vehicle parking occupies the non-motor vehicle lanes and sidewalks road space so that walking and bicycle traffic space is violated causing serious problems and danger in terms of safety. Therefore, in most developed countries and cities curb parking shall be decreased gradually because of the increase of off-street parking lot. Curb parking in Tokyo accounts only for 5% of public parking lot, 4.7% in HK, 22 and 17% in Nanjing and Suzhou respectively. Percentage of curb parking in China relatively remains high due to the insufficiency of off-street parking lot construction and imperfection of transportation system in city [4].

D. Wang (✉) · X. Han · G. Xing
Hohai University Wentian College, Maanshan, China
e-mail: 113177620@qq.com

Fundamental Solutions to Curb Parking Problems

This paper will give a deep analysis from five aspects for the fundamental solutions to curb parking problems.

Perfection of Public Transportation System

Room is existing for perfection in public transportation system in China. Currently, rail transit lines have been built in 19 cities leaving most of other cities behind without perfect rail transit network. Even in big cities such as Beijing, Shanghai, Guangzhou, and Shenzhen, the rail transit transportation with large numbers cannot meet citizens' normal travel demand. In rush hours, congestion is serious and per capita rail length is still insufficient, as shown in Fig. 1.

Rail transportation investment is of high cost with its long-term construction period which takes certain time to build perfect and developed rail transportation network. Ground regular bus costs less investment with immediate effect, but regular bus construction in China relatively lag behind. Figure 2 represents bus ownership for one thousand people. The number reaches 2.4 in Tokyo where owns developed public transportation system, 1.15 in London, 0.84 and 1.01 in Guangzhou and Beijing respectively.

Population density in Beijing is only a quarter of that in Tokyo while the per capita car ownership is half of that in Tokyo. Then, the problem is why congestion exists in Beijing but not in Tokyo where the per capita car ownership is as twice as that in Beijing. Bus travel accounts for 86.6% in Tokyo due to people's habitual choose for public transportation and its perfect bus system. The rail transportation network density in Tokyo is higher than that in Beijing. People in Beijing cannot bear flocks of travelers in rush hours although they do want to take subway to go to work. People have no choice but to drive cars causing increase of vehicles on road which accounts for problems of traffic jam and parking. It is hard to solve curb

Fig. 1 Per capita rail length
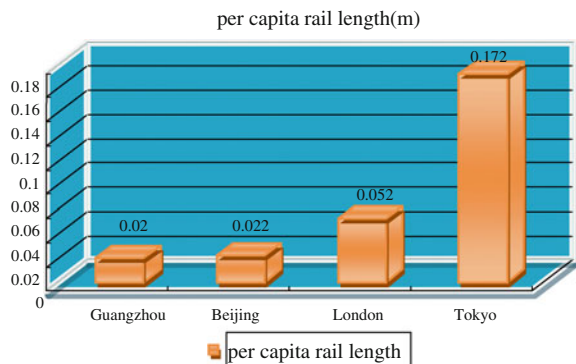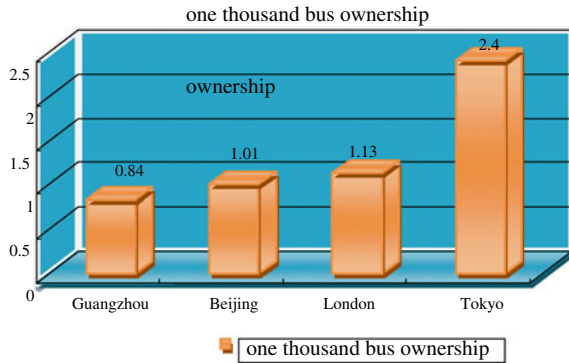
Fig. 2 One thousand bus ownership
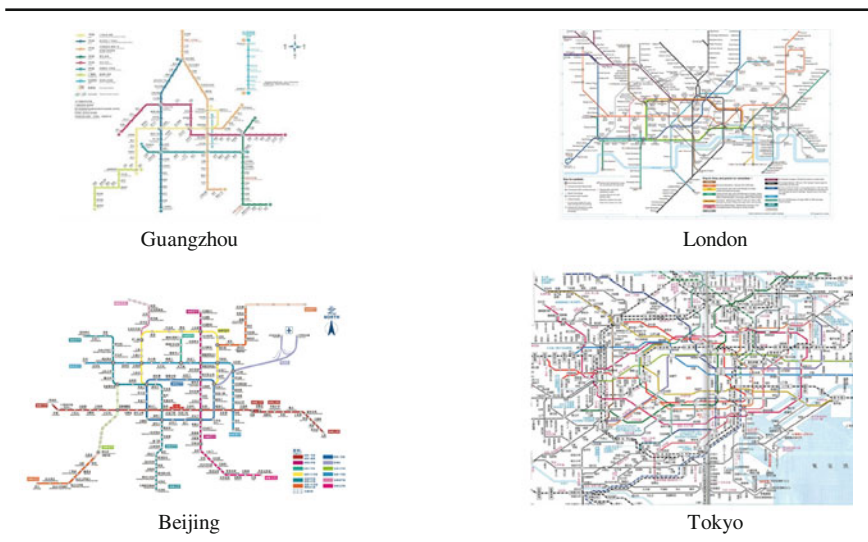
Fig. 3 Comparison of rail transportation network density in four cities

parking problems because of insufficient parking lot. In order to solve curb parking problem ultimately, it is essential to perfect public transportation system (as shown in Fig. 3) [1]. Subway network density in Tokyo in higher than that in Beijing. When perfection of public transportation system in China is realized, percentage of bus in all transportation tools increases. As long as bus traveling percentage in Beijing rises from 46 to 86.6%, curb parking problem is solved naturally benefiting not only the solve of parking issues but also the relief of congestion.

Encouraging the Construction of Mechanical Three-Dimensional Parking Lot

According to statistics, national vehicle ownership reaches 0.225 billion with 0.106 of cars until the end of 2011. National vehicle drivers comes to 0.236 billion. In the last five years, vehicle ownership in China remains high growing rate with an annual increase of 15.91 million and 17.73 million in 2011. According to a survey conducted by the Ministry of Construction, national car parking space gap is extremely high while annual new cars is more than 15 million causing the ratio of parking lot and numbers of cars reaches 1:1.2 which means an annual increase of 2 million parking lot must be provided to meet the parking demand. The rapid development of vehicles further intensifies the contradiction between supply and demand. Insufficiency of parking lot necessarily leads to curb parking problems seriously affecting the traffic capacity. Numbers of parking lot in big cities in China remains low (as shown in Fig. 4). Ratio of parking lot and car ownership reaches 95% in Tokyo, 72% in London, 30 and 31% in Guangzhou and Beijing respectively, suggesting that 70% of vehicles have no room to park in big cities like Guangzhou and Beijing which mainly explains why curb parking problems become serious. Many vehicles even park on roads with no parking lines severely influencing traffic capacity of roads.

Compared with ground parking lot, mechanical three-dimensional parking lot occupies less room with high parking efficiency because it can provide more numbers of parking lots with same area of land. Additionally, many parking lots are plane ones. In this case, when there is little parking lot, it takes long time for drivers finding parking lot. It is possible for novice drivers cannot find space which not only waste time and oil but also cause air pollution. The government shall attempt to encourage construction of mechanical three-dimensional parking lot and adopt some encouraging policies such as decrease of taxation and offer of subsidy in order to meet parking demand in center of cities. From aspect of demand management,

Fig. 4 Ratio of parking lot and car ownership
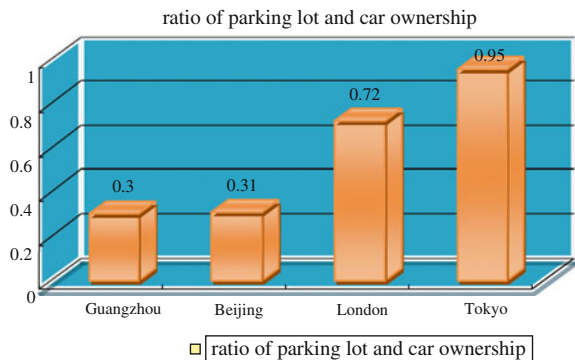
mechanical three-dimensional parking lot can reduce curb parking demand and effectively solve all kinds of problems caused by curb parking as well as its negative influence on dynamic transportation [2].

Prohibition of Raising Parking Fee Blindly

In August 2014, parking fee in Guangzhou grew by 60% and same policy was also adopted in Nanjing and Shenzhen due to experts' opinion that a growth by 60–70% for parking fee can effectively ease the problem of congestion in city. In August, the traffic index changed from 2.5–3.2 to 2.5–7.4 in rush hours meaning that congestion became more serious with 4.2 growth of traffic index peak. Theoretically, use of economic levers can affect the way people travel and higher parking fee indeed reduces the number of vehicles on street so that congestion can be eased and parking demand decreases [3]. However, higher parking fee must bases on advanced public transportation system. Otherwise, blindly raise of parking fee inevitably leads to a series of problems such as the subway-taking rate in rush hours in Beijing reaches 144%, while 95% in Japan. Those people who cannot afford an increase of 60–100% have to take subway or bus to go to work. However, the subway-taking rate in rush hours in big cities such as Beijing, Shanghai, Guangzhou, and Shenzhen has exceeded nominal standard that subway system cannot bear more travel demand anymore. According to survey, travelers hold that taking subway represents loss of dignity. As long as they can afford to buy cars, they prefer to be congested on road than to take subway. Increase of parking fee, which influence little on high income group, make those with medium or low income spend more time to find parking lot. Consequently, parking demand does not decrease but the congestion degree is intensified.

Learning from foreign countries, it is known that parking fee in Singapore, Britain, America and Japan is higher than that in China, but these countries build developed public transportation systems. Public transportation travel percentage in these countries is higher than that in China. 86.6% of travelers in Japan choose public transportation, 70% in London, 60 and 46% in Guangzhou and Beijing respectively. Higher parking fee precisely can effectively solve the problem of congestion and reduce parking demand only when it is based on the perfection of developed public transportation system. Otherwise, it will result in more social problems.

Perfection of Park-and-Ride Facilities

Park-and-ride refers to the set of park-and-ride sites outside of cities such as rail transportation stations, bus stations, and the highway. It aims to provide parking space for free or with low cost for private cars or bicycles supplemented by

Table 1 The user's choice of alternative way to travel to central under the assumptions of cancel park-and-ride facilities

Way	Oxford			York		
	Fri	Sat	A week	Fri	Sat	A week
Cancel travel/change destination	7.2	21.0	10.1	10.5	14.9	11.3
Car	57.0	56.6	56.9	55.5	67.9	58.0
Bus	26.1	15.8	23.9	24.3	9.3	21.3
Others	9.7	6.5	9.2	9.7	8.0	9.3

preferential fee policy for public transportation. It guides passengers transfer into city through public transportation in order to decrease of the use of private cars in center of cities and relief traffic pressure.

Park-and-ride facilities are successfully implemented in Oxford and York. Through periodic rolling survey on the situation before and after implementing park-and-ride and track type continuous monitoring, it finds out that without park-and-ride, about 10% of users will cancel or change their destination suggesting that park-and-ride facilities attract some travel demand. If it is canceled, 56% of travelers in Oxford drive private cars to center of city, 58% in York (shown in Table 1) [5]. Therefore, park-and-ride facilities can effectively decrease the number of cars flocking into the center of cities and control its curb parking demand.

Improvement of Overall Information Intelligent Parking Guidance System

Curb parking problem lies in the fact that limited parking resources cannot meet the rapid growth of parking demand. In order to solve the problem, except increase of parking lot, proper guidance and strengthen of plan, technological measures shall be used to improve parking management standard so that current resources is best used of which is also an efficient method to solve curb parking problem [6].

Parking guidance system includes intelligent guidance system in large scale parking lot and urban parking guidance system, which is based on release of multi-information providing location, number of parking space and state of parking lot in real time. Since the use of "DiDi" and "KuaiDi" software, application of APP in transportation gradually popularized. Bus APP software is used in Beijing, Shanghai, Hangzhou, etc. to look into bus information in real time. APP served in parking has been used in America for a long time, but it only begins to emerge in China. Although the development of software is easy, related hard facilities to match it is still a large systematic engineering. What we need is not only parking information in parking lot but also that nearby. In this case, we can greatly improve the using rate of parking facilities, decrease congestion caused by finding parking lot and the waiting time for parking so that the whole transportation system efficiency is improved and curb illegal parking phenomenon is well solved.

Establishment of overall information intelligent parking guidance system saves much parking time for vehicle travelers and makes the best use of parking space. Curb parking problem can be better solved if vehicle travelers are properly guided instead of given tickets.

Conclusions

In conclusion, fundamental solutions to curb parking problem cannot rely on shouting slogans or sticking tickets. It is a systematic project of large scale. As long as it is well done, both the urban curb parking problem, congestion, and other problems can be solved successfully. Three stages include increase of numbers, perfection of structures, and improvement of service, on which the fundamental solve of curb parking problem depends. It means to perfect parking lot form to increase parking space on basis of current land and enhance the construction of park-and-ride facilities as well as public transportation system. As a result, people's travel structure gets improved and curbs parking demand decreases. It is of tremendous importance to take advantage of overall information intelligent parking guidance system to improve parking service quality. Although it takes certain time to build the systematic project, patience and persistence is required to do it well. Quenching a thirst with poison equals self-killing, only with patience, clean water can be found for human's healthy life.

References

1. Chen, Shidong. 2013. Tightening governing revolution: implementation of the urban parking supply policy. *Journal of Comprehensive Transportation* (2).
2. Chen, Jian. 2007. Research on Parking Strategies in Large Office Area. Master's thesis in Changan University, XiAn.
3. Cai, Huamin. 2001. Research on construction of parking lot and parking management in China. *City Transportation* (4).
4. Fu, Ying. 2011. Research on improvement of center residential parking lot problem in Shanghai-taking old-fashioned department Changfeng square as an example. *Traffic and Transportation*.
5. Yang, Fei. 2005. Park-and-ride effect analysis-taking Oxford and York as examples. *Traffic and Transportation*.
6. Zhu, Chengyao, and Peiliang Qiu. 2013. Analysis of curb parking management for vehicles in center of cities. *Journal of Public Security—Zhejiang Police College* (2).

Empirical Analysis of Hypothetical Bias in Stated-Preference Experiments

Li Tang and Xia Luo

Abstract The stated-preference (SP) survey is widely used to obtain data of choice prediction, especially for new transport service systems. However, the hypothetical nature of SP experiment determines that it is usually taken under a 'what-if' situation. The difference between the 'stated' and 'real' responses is named as the hypothetical bias (HB), which may further lead to great errors of the following model analysis. While HB has been largely confirmed in the area of economics, few studies have proved the existence of HB in transportation planning, not to mention measuring it. This paper presents a new viewpoint to study HB problem in terms of the mode split prediction. An SP survey and a revealed-preference (RP) survey were conducted before and after the opening of a new metro line in Chengdu, China. To reduce requirements on high accurate sampling, a selection–calibration–simulation method is proposed to transfer the individual choice to aggregate market share of alternatives. The multinomial logit (MNL) is selected as the basic model structure in this paper. By calibrating two separate MNL models using SP and RP data, we ran the simulation under a uniform input. The result presents the valid evidence on HB's existence and also indicates the possible direction of deviating it. Furthermore, an improved value of time (VOT) model that can accommodate multiple time and cost variables is proposed. The calculation results of VOT confirm the simulation result and demonstrate how large HB is.

L. Tang
School of Automobile and Transportation, Xihua University,
Chengdu 610039, China

L. Tang (✉) · X. Luo
Jiangsu Province Collaborative Innovation Center, Modern Urban
Traffic Technologies, SiPaiLou Number 2, Nanjing 210096, China
e-mail: tangli@mail.xhu.edu.cn

X. Luo
School of Transportation and Logistics, Southwest Jiaotong University,
Chengdu 610031, China

Introduction

As a major survey method for discrete choice analysis, stated-preference (SP) experiments usually present sampled respondents with a number of different hypothetical situations, each consisting of a finite set of alternatives defined on a number of attribute dimensions. Respondents are required to make choices or state their preference among these situations. The deviation between 'stated' and 'real' choices or valuations is referred to as hypothetical bias (HB). This notion implies that, due to the hypothetical nature of SP experiments, people may behave inconsistently that they do not back up their stated preference with real commitments [1].

Evidence of HB has been largely found and discussed in marketing and environmental area, such as economic valuation for endangered species, public space, and air pollution [2]. In transportation area, HB is usually identified in experiments estimating the value of time (VOT). Although efforts to study HB on marginal willingness to pay (WTP) on travel time have been made in recent years, only a very small number of literature can provide powerful results that the VOT in SP experiments is less than half amount of that in revealed-preference (RP) data [3, 4]. Thus, Hensher [1] mentioned in his 2010 paper that more empirical studies to confirm or deny the growing of evidence on HB in SP experiments are needed.

Another motivation of this paper lies in the wide use of SP experiments for behavior analysis in transportation area. Especially for big cities in developing countries, new transportation modes and innovative traffic management methods, such as building a metro line or conducting congestion pricing, are now extensively introduced or discussed. Since the lack of practice, SP data seems to be the only source for predicting people's reaction toward these changes. If HB really exists in such a big difference as reported in the literature, researchers may need to have a second thought in using SP data directly for model calibration and valuation.

This paper targets on providing an empirical study on HB about its existence and magnitude. We aim to present the following steps to the literature: first, considering that most of the evidences on HB are explored in the aspect of economic analysis, we provide another dimension to represent it—in terms of model prediction results. A selection–calibration–simulation method is proposed to convert individual choice data to aggregated market share of alternatives. Inputting a SP and an RP dataset, respectively, into the selected model structure, the chosen probability of each alternative can be obtained through the calibrated models and be compared in a uniform scenario using RP data as the benchmark. Second, a VOT model is presented to keep consistency with former successful studies in measuring HB. With our empirical travel mode choice data collected in China, we have a valuable chance to discuss whether HB varies with geographic distribution.

Literature Review

Studies on HB were confined largely, but not exclusively, within the field of experimental economics, identified initially as the misrepresentation of preferences [5–7]. In early work of Bohm's paper in 1972, he conducted a test involving six approaches for estimation of the demand of a public good and compared the hypothetical and real values [8]. But it was not caught public's attention until a decade later. In the 1980s and 1990s, the existence of HB is confirmed by testing the overall validity of contingent valuation in a series of experimental HB literature, suggesting that values derived from SP surveys typically exceed actual values. Heberlein and Bishop found that the WTP for hunters buying a hunting permit is 25–33% higher in hypothetical situations than they really take [9]. Fox et al. developed Shogren's CVM-X procedure to measure and correct HB. In their paper, they found customers may overstate their WTP for irradiated/non-irradiated meat in different degrees and suggested calibration factors for hypothetical bids of these two kinds of pork [10, 11]. List and Shogren supported their conclusions by comparing bidding behavior in a hypothetical and actual second-price auction for baseball cards [12]. During that time, most studies point out that the HB in WTP is 2 times larger than the value people take in real situations [13, 14].

In the last few years, attempts of synthesizing the plethora of HB studies are made. Foster et al. present a table summarizing 14 study results which compared real and hypothetical payments for private/public goods. However, due to different elicitation methodologies, they failed to draw a uniform conclusion on calibration factor. Yet, a distinct trend that hypothetical payments exceed at least 1.5–2 times than real ones can be seen [15]. Understanding heterogeneous experimental techniques would cause invaluable magnitude of HB, List and Gallet examine data from 29 experimental studies using a meta-analysis in 2001. Their findings suggest that the overstatement of preferences is 3 times larger in hypothetical settings [16]. Murphy et al. conducted another meta-analysis of HB in 2005, more specifically using 28 stated-preference studies with same mechanism for eliciting WTP values. The median ratio of hypothetical to actual value is only 1.35 in their report, explained as a choice-based elicitation mechanism is important in reducing HB [2]. Harrison and Rutstrőm later provided a detailed and systematic analysis on HB from a wider literature. But they avoid drawing any board conclusion, only claiming values estimated in hypothetical situations usually exceed real ones [17].

Economic valuation in transportation behavior area is largely narrowed in estimating travelers' VOT. Contrasted with the general conclusion drawing from marketing area, it is interesting to investigate in most research that stated VOT is smaller than real ones. One influential paper by Brownstone and Small concludes, based on a toll road context in California, that commuters' real VOT are at least twice as large as SP results. Although they admit part of the bias lies in the methodologies of the survey techniques, they strongly suggest that using SP data

directly will undervalue the benefit from travel time saving [3]. Following their research, Isacsson support Brownstone and Small's conclusion in his 2007 paper, in the context of trading time with money, that the marginal WTP based on the hypothetical experiment was almost 50% lower at the mean than the real ones. Instead of using a conventional random utility model, he tests the differences between real and hypothetical VOT by directly comparing cumulative distribution functions of VOT and fitting parameter distributions [4].

However, it is too early to optimistically report that HB does exist in transportation area as in other fields. Beside Brownstone and Isacsson, there are other researchers who also trend to reject its existence. Smith and Mansfield find no significant differences between people's choices with real and hypothetical offers in a field test regarding VOT [18]. In the survey, respondents are given real and hypothetical offers to participate in a future survey in exchange of money. But it does not specify the exact time for the future survey. So the scheduling constraint may disturb people's judgment, explained by Isacsson. In 2010, Hensher draw on a number of datasets over the last 20 years and calculate the VOT for each RP and SP choice set in separate mixed logit error component models. The ratio RP/SP is not significantly different from 1.0 in all the studies and hence he fails to reject the null hypothesis of no evidence of HB [1]. One possible explanation for this is that traditional RP estimate is not a benchmark compared to real market observation. Though Brownstone and Small use a usual RP model, the data is obtained from a sample of travelers' actually observed choosing between a variable priced tolled lane and a free lane. The attributes are measured by external procedures so that the levels of time and costs actually experienced for both alternatives in the choice set are not subject to the usual concerns associated with asking individuals. But this assumption has not been verified due to lack of following research.

In summary, although HB has been found and confirmed in marketing and environmental areas, the uniform conclusion about its existence in transportation cannot be reached from these contradictory studies. We also notice that all the existing research focused on economic valuation of travel time; investigations about whether HB exists in prediction result are limited. Therefore, a specific and detailed analysis of the evidence on HB, and how it makes effect on prediction and VOT are highly needed.

Data Spectrum for Empirical Analysis

Notation

For convenience, the following notations are listed in Table 1 for generation of experimental design and formulation of travel mode choice model.

Table 1 Notation specification

Notation	Description
TC	Travel cost (RMB)
RT	Time to reach at bus or subway station (min)
WT	Wait time in bus or subway station (min)
TT	Travel time (min)
AT	Time to arrive at final destination (min)
PT	Time for parking (min)
Gndr	Gender of respondent: male 1; female 0
Car	Car owning condition of respondents: at least one 1; none 0
Age	Age of respondents: under 18 1; 18–24 2; 25–30 3; 31–35 4; 36–40 5; 41–45 6; 45–50 7; 50–60 8; above 60 9
Incm	Monthly income of respondents: under ¥2000 1; ¥2001–4000 2; ¥4001–6000 3; ¥6001–8000 4; above ¥8000 5
Aim	Trip purpose: office 1; school 2; visiting 3; recreation 4; otherwise 5
ASCBUS	The attribute-specific constant for alternative bus
ASCTAXI	The attribute-specific constant for alternative taxi
ASCCAR	The attribute-specific constant for alternative car

Experimental Design

The travel mode choice data used for this analysis is collected on a corridor connecting suburb and urban areas in Chengdu, China. Respondents are randomly sampled along the corridor. Four alternatives are involved in the study: car, taxi, bus, and subway. Among them, the former three are long-existing transport modes and the subway, No. 2 West Extended Line, is newly introduced. The No. 2 West Extended Line is officially operated on June 8, 2013. The opening of this subway line is expected to provide a new convenient public transit service and help mitigate the heavy traffic congestion along this corridor. Also, it is a great opportunity for researchers to investigate how travel mode choice behavior changes with the evolution of city traffic structure.

The survey experiment is conducted in two stages: first, an SP survey is taken from May 1 to June 4, 2013. Respondents are required to state their travel mode choice among the above four alternatives in different choice situations. In this stage, the subway is still under construction, and thus it is regarded as a 'hypothetical' alternative in the choice set. Orthogonal design method is applied to the SP experiment for an independent determination of each attribute's influence upon the observed choices. The attribute levels are given in Table 2 based on previous study results as well as to preserve realistic estimates for the private and public transport alternatives. To generate an orthogonal design under such an attribute level setting,

Table 2 Attribute levels and values for SP survey

Alternative	Attribute	Attribute value		
Bus	*TC* (RMB)	2	4	
	RT (min)	5	10	
	WT (min)	5	10	
	TT (min)	40	50	60
	AT (min)	5	10	
Taxi	*TC* (RMB)	25	30	
	WT (min)	5		
	TT (min)	20	30	
Car	*TC* (RMB)	20		
	TT (min)	20	30	
	PT (min)	5	10	
Subway	*TC* (RMB)	2	3	
	RT (min)	10	20	
	WT (min)	3		
	TT (min)	15	20	
	AT (min)	10		

the minimum number of choice situations is 16, which is obviously too large for a single person. Thus, a block variable is introduced to divide the design into smaller sections (i.e., here we block the design into four parts so each respondent is faced with four situations). In the second stage, from April 9 to 18, 2014, an RP survey was conducted to obtain the actual travel mode choices on the experiment corridor, along with the real value of attributes [19]. Also, the social-demographic characteristic (SDC) data including gender, car ownership, age, income, and trip purpose is collected in both surveys.

Data Description

360 respondents and 1440 observations are obtained from SP survey. 910 respondents and 910 observations are obtained from RP survey. Since the purpose of this paper is to discuss whether HB exists between RP and SP, it is more important to exam these two sampled datasets rather than compare them with the population. Figure 1 shows the comparison diagrams of SDC variables between SP and RP surveys. It can be seen that the SDC distribution of SP survey generally matches that of RP survey. It should be noticed that more travelers choose bus in reality (28.4% vs. 40.22%), although in SP quite a few people stated to choose subway (43.75% vs. 35.38%). This would be a preliminary evidence of HB.

Gender

	male	female
SP	58.33%	41.67%
RP	51.87%	48.13%

Car Ownership

	yes	no
SP	34.44%	65.56%
RP	29.56%	70.44%

Age

	under 18	18-24	25-30	31-35	36-40	41-45	45-50	50-60	above
SP	5%	30.83%	26.94%	16.39%	8.06%	5%	3.06%	3.61%	1.11%
RP	3.52%	30.22%	27.14%	15.05%	8.79%	4.84%	3.08%	4.51%	2.86%

Income

	under 2000	2000-4000	4000-6000	6000-8000	above 8000
SP	25.00%	46.11%	18.89%	6.94%	3.06%
RP	20.55%	43.19%	23.96%	8.68%	3.63%

Travel Mode Choice

	bus	taxi	car	subway
SP	28.40%	4.24%	23.61%	43.75%
RP	40.22%	6.81%	17.58%	35.38%

Fig. 1 Comparison diagrams of social-demographic characteristic

Identifying HB in Market Share Prediction

There is little literature about whether HB exists in travel mode choice prediction results. There are two possible causes for this: first, the aim of the research is not finding HB but the time value or some other goals. HB in VOT is the incidental discovery during the study process [3], or some research is the replication of others, and thus the researchers should keep the objective in consistence with former ones [4]. Second and the most important, the research object is usually subjected to the data source, especially in the field of empirical study. In literature, most studies on HB in travel behavior analysis are involved with route choice in commuting/ non-commuting trip (i.e., choosing between a toll road and a free road or car-pool lane and regular lanes), and the travel mode choice dataset is merely seen. The reason for this may lie in the fact that in most developed countries, the constructions of urban traffic instruments (e.g., subway, light rail, freeway) have been finished, which implies that the market share of travel mode is stable. Unlike that, the urban traffic structure is under great change in developing countries and new travel mode is extensively introduced into big cities. Thus, the empirical data we use in this paper creates a valuable chance to study how HB works on travel mode choice prediction results.

Since HB is a type of systematic bias instead of sampling bias, the most ideal pattern of discussing HB should compare the SP and RP answers from a same group of respondents. However, it is almost impossible to find such perfect data in reality, especially when facing with big-scale sampling. Thus, an approach that can convert these individual choices to aggregated results is required. Here, we propose a selection–calibration–simulation method based on our SP and RP datasets, as shown in Fig. 2. The basic idea of this approach consists of three stages: (1) Selecting primary analysis model structure according to the characteristic of the choice set. Then using SP data as the input, determine the utility function and the error term distribution. With the SP data, models can be calibrated and the performance of each model structure can be obtained. Choose an appropriate model structure in terms of overall significance and model fit. (2) Importing both SP and RP data separately into the selected model structure, get the calibrated models base on these two datasets. (3) To get comparable aggregated predict results from the calibrated models, a uniform scenario should be settled and reentered into each model. We call this process as simulation. Then the aggregated market share of each alternative can be obtained for the final comparison between SP and RP data.

Stage 1: Selection

To reduce the interference from model itself during the data processing, it is necessary to determine a proper model structure at the initial stage. The empirical dataset in this paper involves with four alternatives: bus, taxi, car, and subway.
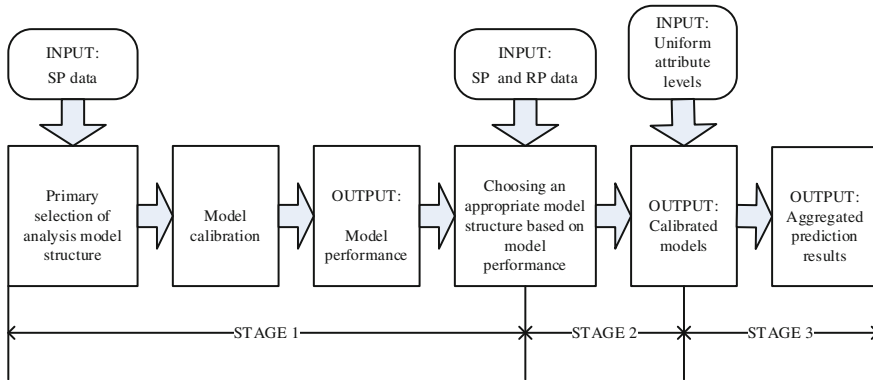
Fig. 2 Framework of identifying HB in prediction results

Considering their characteristics, MNL and NL models are prepared for the primary selection.

To build a MNL model, assume that an individual n is faced with a choice set J, the utility function for that individual choosing alternative $i, i \in J$ can be expressed as

$$U_{in} = V_{in} + \varepsilon_{in}, \tag{1}$$

where V_{in} denotes the deterministic part of utility that can be observed by the analyzers and it can be formulated as the liner addition of explanatory variables (i.e., including alternative-related attribute variable and social-demographic characteristic of individual listed in Tables 1 and 2) with corresponding parameters [20]. ε_{in} denotes the random error term which is assumed to be independently and identically extreme value type I distributed. Thus the probability of individual n choosing alternative i among the choice set J can be expressed as

$$P_{in} = \frac{\exp(V_{in})}{\sum_{j \in J} \exp(V_{jn})}. \tag{2}$$

Using notation adopted from Bliemer [21], a two-level nested logit tree is built for analysis. The upper level consists of two branches: Prvt short for private and Pblc short for public. Each branch contains two elemental alternatives: car and taxi in the private branch; bus and subway in the public branch.

Within the NL, each branch and elemental alternative will have an associated scale parameter. Here, let λ_m denote the scale parameter of branch m, and let $\mu_{j|m}$ denote the scale number of elemental alternative j within branch m. Let J_m denote the set of all elemental alternatives belonging to branch m. By definition, all scale parameters of the elemental alternatives in this set J_m have the same scale

parameter. That is, $\mu_{j|m} = \bar{\mu}_m$ for all $j \in J_m$ for some value $\bar{\mu}_m$. Thus, the utility for an individual n choosing an elemental alternative i in branch m can be expressed as

$$U_{in|m} = V_{in|m} + \varepsilon_{in|m}, \tag{3}$$

where $V_{in|m}$ denotes the observed component of utility and $\varepsilon_{in|m}$ denotes the unobserved random error, which is also identically and independently extreme type I distributed. The probability of choosing elemental alternative i given that branch m was chosen can therefore be seen as a simple MNL model and expressed as

$$P_{in|m} = \frac{\exp(V_{in|m})}{\sum_{j \in J_m} \exp(V_{jn|m})}. \tag{4}$$

The observed utility for branch m can be expressed as the product of inclusive value (IV) variable (i.e., the expected maximum utility index to link branch m and its elemental alternatives, which is equal to the nature logarithm of the denominator of the MNL model associated with the elemental alternatives) and the branch scale parameter:

$$V_{mn} = \frac{\lambda_m}{\bar{\mu}_m} \log\left(\sum_{j \in J_m} \exp(V_{jn|m})\right). \tag{5}$$

The probability of choosing branch m is given by

$$P_{mn} = \frac{\exp(V_{mn})}{\sum_{m=1}^{M} \exp(V_{mn})}. \tag{6}$$

Combining Eqs. (4)–(6), an individual n choosing elemental alternative i can therefore be expressed as

$$P_{in} = P_{mn}P_{in|m} = \frac{\left(\sum_{i \in J_m} \exp\left(V_{in|m}\right)\right)^{\lambda_m/\bar{\mu}_m}}{\sum_{m=1}^{M} \left(\sum_{j \in J_m} \exp\left(V_{jn|m}\right)\right)^{\lambda_m/\bar{\mu}_m}} \frac{\exp(V_{in|m})}{\sum_{j \in J_m} \exp(V_{jn|m})}. \tag{7}$$

To analyze above models with the empirical dataset, we use the professional software called NLOGIT. It is a complete economics and statistics package for estimation and simulation of multinomial discrete choice models. Though the model structures of MNL and NL are different, NLOGIT uses same criteria to evaluate model performance. Seeing from the output, the probability values (p value) of both MNL and NL models equal to zero, which implies that both of them have an overall model significance. To compare the MNL and NL models, the log-likelihood (LL) ratio test is conducted. NLOGIT is able to generate the value of LL function in the output. For MNL model, the LL function estimate is -1558.581

and for NL model this value equals to -1359.279. With the values of LL function, the performance of MNL and NL can be compared in terms of model fit with the R^2 index, which can be expressed as

$$R^2 = 1 - \frac{\text{LL}_{\text{NL}}}{\text{LL}_{\text{MNL}}}. \tag{8}$$

Substituting the values from NL model output (-1359.279) and MNL model output (-1558.581) in Eq. (8), the R^2 value is calculated as 0.128. This result indicates that there is no significant improvement of changing model structure from MNL to NL since a value of 0.3 usually represents a decent model fit for a discrete choice model. Thus, MNL model is chosen as the basic model structure for the following analysis.

Stage 2: Calibration

The MNL model can be calibrated using maximum likelihood estimation (MLE). Using notation adapted from Ben-Akiva [22], let N denote the total observations and define

$$y_{jn} = \begin{cases} 1 & \text{if observation } n \text{ choose alternative } j \\ 0 & \text{otherwise} \end{cases}. \tag{9}$$

Assuming a single respondent, the log-likelihood function can be expressed as

$$L = \sum_{n=1}^{N} \sum_{j=1}^{J} y_{jn} \log P_{jn}. \tag{10}$$

To obtain the maximum estimation of LL function, set the first deviation of L with respect to the parameters equal to zero, expressed as

$$\frac{\partial L}{\partial \beta_k} = = \sum_{n=1}^{N} \sum_{j=1}^{J} \left[y_{jn} - P_{jn} \right] x_{jnk}, \tag{11}$$

where β_k denotes the parameter of attribute $k, k \in K$ and x_{jnk} denotes the value of it. The estimation for parameters in the proposed MNL model using SP and RP data is shown in Table 3.

The probability values (known as p value) are provided as the criterion for determining significant parameters. Assuming a 95% confidence level, p value less than the determined level of alpha which equals 0.05 may reject the hypothesis that the parameter equals to zero and suggests that the explanatory variable is statistically significant. These significant corresponding coefficients are marked with '*' in

Table 3 Calibration result using SP and RP data

Dataset variable	SP data		RP data	
	Coefficient	p value	Coefficient	p value
TC1	−0.181*	0.001	−0.298*	0.000
TC2	−0.021*	0.049	−0.021*	0.024
RT	−0.112*	0.000	−0.022	0.051
WT	0.030	0.218	0.003	0.823
TT1	−0.014*	0.049	−0.016*	0.006
TT2	−0.037*	0.001	−0.024*	0.048
AT	−0.022	0.370	−0.052*	0.000
PT	0.014	0.594	−0.070*	0.012
Gndr	−0.152	0.206	0.330*	0.024
Car	1.062*	0.000	0.789*	0.000
Age	0.100*	0.005	0.068*	0.095
Incm	0.395*	0.000	0.129	0.136
Aim	0.177	1.000	−0.034	0.486
ASCBUS	−2.373*	0.000	−0.717*	0.020
ASCTAXI	−4.921*	0.001	−3.120*	0.000
ASCCAR	−3.296*	0.006	−2.201*	0.000

Table 3. It is notable that we separate *TC* as *TC1* and *TC2*, *TT* as *TT1* and *TT2* for the convenience to calculate VOT of each alternative in Section "HB Measurement by VOT", where *TC1* denotes the parameter of travel cost using bus and subway and *TC2* denotes that of travel cost using taxi and car. *TT1* denotes the parameter of travel time using bus and subway and *TT2* denotes that of travel time using taxi and car. As shown in Table 3, the parameters for most critical explanatory variables are statistically significant. Inconsistencies can be seen in *RT*, *AT,* and *PT* for alternative-related attributes and *Gndr* and *Incm* for SDC. It suggests that in different situations (i.e., hypothetical or real) the impact of each attribute for respondents may differ, or it is simply because of the sampling problem of these two datasets.

Stage 3: Simulation

The capability of simulation is to test how changes in attributes and SDC impact upon the choice probabilities for each of the alternatives. For keeping the simulation close to reality, the attribute mean value from RP data is used as the input. Then the market share can be recalculated with the calibrated model based on SP and RP data, respectively.

It is noteworthy that there is no change on the values of SDC, so that the hypothetical/real scenario remains to be a specific individual. In other words, to a particular decision-maker, the condition of making choice (i.e., stated or revealed)

Table 4 Simulation result

Choice	SP scenario		RP scenario		RP–SP	
	Percentage (%)	Number	Percentage (%)	Number	Percentage change (%)	Number change
Bus	28.116	405	41.127	374	13.01	−31
Taxi	3.279	47	7.065	64	3.786	17
Car	21.196	305	17.536	160	−3.66	−145
Subway	47.408	683	34.272	312	−13.136	−371

has not changed and the only difference is the alternative-related choice situation. Aggregating individuals' choices, the simulation result is demonstrated in Table 4.

As shown in Table 4, there is a significant difference between the prediction results for each alternative when using models calibrated in SP/RP situation with the same input. In general, for the type of travel modes (i.e., public or private), people tend to behave consistently (75.524% vs. 75.399% in bus and subway, 24.475% vs. 24.601% in taxi and car). However, for a specific travel mode, an apparent increase of choices on bus can be observed while an obvious decrease of choice on subway is found. It indicates that in hypothetical situations, people may overestimate their preference on new type of transit and neglect their reliance or habit on the traditional ones. Overall, this result is not only a powerful evidence that proves HB does exist, but also suggests the pattern of the bias and its possible size.

HB Measurement by VOT

Improvement on VOT Model

In the limited literature which has successfully proved the existence of HB, travelers' VOT is usually chosen as the measure and comparison criteria, although in Section "Identifying HB in Market Share Prediction", we have identified HB in terms of market share and shown the significant differences between SP and RP in Table 4. To further confirm our result, VOT of each alternative is calculated using an improved VOT model in this section.

Ever since Beesley laid the foundation of valuing time spent in traveling in 1960s [23], the measurement of VOT has been developed in various formats such as allowing heterogeneity in decision-makers [24] or investigating the distribution of VOT using generalized method of moments [25]. However, using these methods requires high-quality data. Complex modeling and computation work cannot also be avoided. For MNL model, the utility function is linear. Thus VOT is travel time measured in monetary terms, expressed as

$$\text{VOT} = \frac{\beta_{\text{time}}}{\beta_{\text{cst}}} \times 60, \tag{12}$$

where β_{time} denotes the parameter estimate for the travel time-related explanatory variable and β_{cst} denotes the parameter estimate for cost. The VOT is multiplied by 60 so that it can be described as RMB/h. With the above equation, VOT for each alternative is able to be calculated if time or cost variable is alternative specific. However, this conventional linear model cannot be applied to utility functions with multiple time or cost explanatory variables. Here, an improved VOT model is presented to solve this problem. Scale parameters are added to the first deviation of utility functions with time and cost. Thus the VOT of alternative i can be expressed as

$$\text{VOT}_i = \frac{\sum_{k=1}^{K} \lambda_k \frac{\partial U_i}{\partial T_k}}{\sum_{m=1}^{M} \gamma_m \frac{\partial U_i}{\partial C_m}} \times 60, \tag{13}$$

where U_i represents the utility of alternative i, T_k represents the kth time-related variable, and C_m represents the mth cost-related variable. λ_k and γ_m are the scale parameters for the corresponding deviations. For our empirical model, the VOT function is still linear. Thus, λ_k and γ_m can be expressed as

$$\lambda_k = \frac{\beta_{T_k}}{\sum_{k=1}^{K} \beta_{T_k}} \tag{14}$$

$$\gamma_m = \frac{\beta_{C_m}}{\sum_{m=1}^{M} \beta_{C_m}}, \tag{15}$$

where β_{T_k} denotes the parameter estimate for T_k and β_{C_m} denotes the parameter estimate for C_m.

Result Discussion

Combining Eqs. (13)–(15), the VOT of each alternative in SP- and RP-based MNL model can be obtained, as shown in Table 5.

Table 5 VOT of each alternative in SP and RP data

Choice	VOT in SP data (RMB/h)	VOT in RP data (RMB/h)	VOT ratio-SP/RP
Bus	33	9	3.7
Taxi	106	68	1.6
Car	106	166	0.6
Subway	33	9	3.7

As we can observe from Table 5, the SP results imply a significant difference from the RP results. For car users, the VOT from SP responses is about half that of RP, which is consistent with other research results mentioned earlier. An interesting phenomenon can be observed that for bus, taxi, and subway users, their SP estimates are much lower than the RP estimates. While it is complicated to fully explain this contradictory result, two possibilities are considered for future analysis. First, though a same model structure is chosen for calibration, the separate estimation results using SP and RP data do not perfectly match. The time-related parameters in SP are *TT1*, *TT2,* and *RT,* while that in RP are *TT1*, *TT2*, *AT,* and *PT*. These parameter estimates will largely affect the calculation of VOT with our proposed model. Second, the sign of HB in VOT may vary from alternative to alternative. This explanation connects how people value time with the specific travel mode they take or preferred to take. For example, car drivers may intend to spend less money but are forced to pay more by a scheduling constraint. To confirm both of these two considerations, more datasets and more sophisticated models are required. Nevertheless, these results continue to favor the conclusion in Section "Identifying HB in Market Share Prediction" that HB exists in SP and RP studies.

Conclusions and Future Work

The SP experiment has been long regarded as an effective method for behavior analysis, especially for newly introduced alternatives. However, analysts merely suspect how 'real' these responses are from the hypothetical situations and what the possible influences it will make to the whole methodology. This paper presents a comprehensive and systematic discussion on HB using two precious datasets collected in Chengdu, China.

Considering that most of the former studies on HB are the appendage from other objectives, a new aspect to find HB in travel mode choice is provided in this paper. Knowing that data is the greatest challenge in conducting such a kind of research, a selection–calibration–simulation method is proposed to lower the requirements on sampling accuracy. Three stages are designed to convert individual's preference to aggregated prediction results. Under the same selected model structure, two separate models are calibrated using SP and RP data and the simulation is conducted with the same input. The result of simulation reveals that people do behave inconsistently between the SP experiment and their real acts, which provides a valid evidence of HB to the very limited literature. Furthermore, the results also indicate that public transit users may have much stronger reliance on their old travel habit than expected, which would help transport planners improve their evaluation of public acceptance. In addition, to measure the size of HB, an improved VOT model is proposed. How people value different times when using a specific alternative is calculated. A significant difference on VOT is also found. The result further confirms the existence of HB.

Future work of this study includes the following aspects. First, whereas the main goal is to explore HB in SP surveys, only some basic attributes are included in the utility function of this paper. For better model fitting, more advanced model structure that allows heterogeneity such as mixed logit or hybrid logit should be applied. Second, the result of the VOT model implies that HB may not always be negative (i.e., the SP estimate is not always lower than that in RP). More empirical work is needed to investigate this deviation in different alternatives.

References

1. Hensher, D.A. 2010. hypothetical bias, choice experiments and willingness to pay. *Transportation Research Part B* 44 (6): 735–752.
2. Murphy, J.J., P.G. Allen, T.H. Stevens, and D. Weatherhead. 2005. A meta-analysis of hypothetical bias in stated preference valuation. *Environmental and Resource Economics* 30: 313–325.
3. Brownstone, D., and K. Small. 2005. Valuing time and reliability assessing the evidence from road pricing demonstrations. *Transportation Research Part A* 39 (4): 279–293.
4. Isacsson, G. 2007. *The Tradeoff Between Time and Money: Is There a Difference Between Real and Hypothetical Choices?* Borlange, Sweden: Swedish National Road and Transport Research Institute.
5. Dreze, J.H., and D. de la Vallee Poussin. 1971. A tâtonnement process for public goods. *The Review of Economic Studies* 38 (2): 133–150.
6. Hylland, A., and R. Zeckhauser. 1979. The efficient allocation of individuals to positions. *Journal of Political Economy* 87 (2): 293–314.
7. Roberts, J. 1979. Incentives in planning procedures for the provision of public goods. *The Review of Economic Studies* 46 (2): 283–292.
8. Bohm, P. 1972. Estimating the demand for public goods: An experiment. *European Economic Review* 3: 111–130.
9. Heberlein, T.A., and R.C. Bishop. 1986. Assessing the validity of contingent valuation: Three field experiments. *The Science of the Total Environment* 56: 99–107.
10. Shogren, J.F. 1990. The impact of self-protection and self-insurance on individual response to risk. *Journal of Risk and Uncertainty* 3 (2): 191–204.
11. Fox, J.A., J.F. Shogren, D.J. Hayes, and J.B. Kliebenstein. 1998. CVM-X: Calibrating contingent values with experimental auction markets. *American Journal of Agricultural Economics* 80: 455–465.
12. List, J.A., and J.F. Shogren. 1998. Calibration of the difference between actual and hypothetical valuations in a field experiment. *Journal of Economic Behavior & Organization* 37: 193–205.
13. National Oceanic and Atmospheric Administration. 1994. Natural resource damage assessment: Proposed rules. *Federal Register* 4 (59): 23098–23111.
14. National Oceanic and Atmospheric Administration. 1996. Natural resource damage assessment: Final rules. *Federal Register* 5(61): 439 pp.

15. Foster, V., I.J. Bateman, and D. Harley. 1997. Real and hypothetical willingness to pay for environmental preservation: A non-experimental comparison. *Journal of Agricultural Economics* 48(1–3): 123–137.
16. List, J.A., and C.A. Gallet. 2001. What experimental protocol influence disparities between actual and hypothetical stated values? *Environmental & Resource Economics* 20 (3): 241–254.
17. Harrison, G.W., and E.E. Rutström. 2008. Experimental evidence on the existence of hypothetical bias in value elicitation methods. *Handbook of Experimental Economics Results* 1: 752–767.
18. Smith, V.K., and C. Mansfield. 1998. Buying time: Real and hypothetical offers. *Journal of Environmental Economics and Management* 36: 209–224.
19. Mao, L., and L. Tang. 2014. Analysis of travel characteristics of metro extension line. A case study in Chengdu China. ICLTM 2014.
20. Train, K.E. 2003. *Discrete Choice Methods with Simulations*. Cambridge UK: Cambridge University Press.
21. Bliemer, M.C.J., J.M. Rose, and D.A. Henser. 2009. Efficient stated choice experiments for estimating nested logit models. *Transportation Research B* 43 (1): 19–35.
22. Ben-Akiva, M., and S.R. Lerman. 1985. *Discrete Choice Analysis: Theory and Application to Travel Demand*. MIT Press.
23. Foster, C.D., and M.E. Beesley. 1963. Estimating the social benefit of constructing an underground railway in London. *Journal of the Royal Statistical Society. Series A (General)* 126(1): 46–93.
24. Small, K.A., C. Winston, and J. Yan. 2005. Uncovering the distribution of motorists' preferences for travel time and reliability. *Econometrica* 73 (4): 1367–1382.
25. Fosgerau, M. 2006. Investigating the distribution of the value of travel time savings. *Transportation Research Part B* 40: 688–707.

Real-Time Density-Based on-Ramp Metering Algorithm Considering Multi-Lane of Mainstream

Li Tang, Xunfei Gao, Pengfei Zhai and Xia Luo

Abstract On-ramp metering is regarded as one of most effective control methods of balancing mainstream traffic flow and releasing congestion on expressway in recent years. The current control methods suffer from several shortcomings, such as lack of consideration of on-ramp queue or lane-change behavior on multi-lane mainstream. This paper optimizes the real-time density-based on-ramp metering algorithm by taking multi-lane traffic flow character into consideration. The error function representing lane difference is defined as the objective, and real-time density considering lane change is used as the control parameter of calculating the metering rate. The micro-simulation is used for testing the performance of the multi-lane real-time density-based on-ramp metering (MRD-RM) model using the field data collected in Chengdu. The simulation result shows that MRD-RM model outperforms ALINEA and non-control in terms of reducing average queue length as well as keeping traffic flow on mainstream close to capacity.

Keywords Intelligent transportation systems (ITSs) · On-ramp metering · Multi-lane · Real-time density · Lane-changing rate

L. Tang (✉)
School of Automobile and Transportation, Xihua University,
Chengdu 610039, China
e-mail: tangli@mail.xhu.edu.cn

L. Tang · X. Luo
Jiangsu Province Collaborative Innovation Center, Modern Urban Traffic
Technologies, SiPaiLou Number 2, Nanjing 210096, China

X. Gao · P. Zhai · X. Luo
School of Transportation and Logistics, Southwest Jiaotong University,
Chengdu 610031, China

Introduction

Congestion on urban expressway will cause traffic delay, air pollution, and safety issues [1]. It not only influences the operation of mainstream, but also brings congestion to connected roads when on-ramp long queue happens. All recent traffic flow theories agree that even the traffic demand is larger than capacity, the downstream traffic volume cannot reach the capacity due to the confluence conflict in the merge area between the upstream traffic flow and the merging traffic flow of on-ramp under non-control condition [2]. Thus, in order to maximize the downstream flow, it is necessary to implement intelligent control on the mainstream and its connected ramp.

By far, on-ramp metering is regarded as the most effective and widely used method to release the congestion of urban expressway. By adjusting the on-ramp flow, the downstream flow can be maximized (but not exceeding the capacity). To date, plenty of empirical studies of on-ramp metering as well as simulations have been conducted in many countries, such as USA, French, and the Netherlands. The field experience and simulation results suggest that using on-ramp metering can be significantly helpful in the aspects of increasing the mainstream speed, decreasing the total travel time, and reducing fuel consuming and air pollution.

The on-ramp metering algorithm can be divided into pre-timed metering and responsive metering based on the involvement of real-time information. Among them, ALINEA (Asserv-issement Linéaire d'Entrée Autoroutière-ALINEA) is the most cited control strategy. Since Papageorgiou presented its original algorithm in 1991, a lot of researchers contribute to optimizing the ALINEA-based algorithm from different standpoints. However, although the downstream occupancy o_{out} can be stabilized around the expected occupancy $\hat{o}$ when using ALINEA, all these methods suffer from one shortage that the output of metering rate is relatively small caused by lacking consideration of on-ramp queue. Therefore, when mainstream traffic flow gets close to road capacity, it is highly possible that an overaccumulation of vehicles of on-ramp will be seen. Under this situation, real-time density is proposed as the optimization objective by some other scholars. In this way, the mainstream traffic flow density can be kept close to the expected value while the queueing vehicles on ramp reduced as much as possible [3]. However, the current method does not give enough attention to the influence of lane-change behavior on multi-lane mainstream in a real situation.

This paper further optimizes the real-time density-based on-ramp metering algorithm by taking multi-lane traffic flow character into consideration. The error function representing lane difference is defined as the objective, and real-time density considering lane change is used as the control parameter of calculating the metering rate. The micro-simulation is used for testing the performance of the multi-lane real-time density-based on-ramp metering (MRD-RM) model using the field data collected in Chengdu.

Literature Review

The study on ramp metering strategies and algorithms began in 1960s; based on the level of response to traffic, there are two types of ramp metering control: pre-timed metering and responsive metering; in view of the area of control, responsive metering can be divided into two categories: local (isolated) control and coordinated (area-wide) control [4, 5].

Responsive metering makes use of real-time traffic information detection technique to obtain volume, speed, and occupancy (density) data needed. Responsive local ramp strategy is one branch of responsive metering, which uses specific algorithm in local ramp based on real-time traffic data collection. DC, an early simple dynamic control way, defined metering rate as different values between downstream freeway capacity and upstream freeway flow; at the same time, the downstream critical occupancy is used as judgment criteria; if the downstream occupancy becomes overcritical, the metering rate is reduced to the minimum flow [6]. Based on the same philosophy, OCC uses upstream occupancy detection value to estimate upstream flow and then confirms the metering rate [6]. Above two algorithms can be characterized to be open-loop, both algorithms are quiet sensitive to the various further non-measurable disturbances, and system stability is poor. An alternative ALINEA is a close-loop ramp metering strategy, and the problems of sensitivity to disturbances of open-loop have been successfully resolved. ALINEA is a control strategy developed from classical control theory represented by PID, the metering rate is codetermined by the last metering rate and the real-time downstream occupancy. Obviously, this feedback law is simpler than other local metering strategy; moreover, ALINEA reacts smoothly even to slight errors and stabilizing the traffic flow around the desire value, and thus, congestion may be prevented [7]. Nevertheless, the applications of ALINEA revealed series issues and needs that are not covered by ALINEA, so a number of modifications and extensions are developed that enable some new aspects to be taken into account, such as METALINEA [8, 9], UP-ALINEA [2, 10], X-ALINEA/Q [10], AD-ALINEA, AU-ALINEA [11], and PI-ALINEA [12, 13].

The study on multi-lane traffic flow accords with the need of rapid development of city from the point of macroscopic view, and in microscopic scale, the characteristics of multi-lane traffic flow are be better understand. The research on multi-lane traffic flow is just starting, but a number of achievements have been obtained. Based on the Nagel–Schreckenberg (NS) model, a three lane cellular automate traffic flow model is proposed to describe the highway traffic under the periodic conditions. In this model, all cars obey the same evolution rules and are allowed to change lane probabilities, and the result shows that the traffic volume is higher than that of the NS model and that reducing the lane-changing rate is helpful to safe drive [14]. Based on the NS model, a mixing traffic flow model is also presented for a two-lane traffic system. In this mode, the stochastic noise in NS

model is replaced by accelerating probability and braking probability. The result proves that the lane-changing probability, the mixing rate and acceleration, and the braking probability have a great influence on the mixing traffic flow, and the slow car has a decisive action on the mixing traffic [15]. A bilevel hierarchical control model of on-ramp metering is conducted, which is based on the dynamic traffic flow model in multi-lane freeway; after the theoretical description, a simple example is described, in where the traffic flow in each lane tends to stability and the destiny tends to a constant quantity [16, 17]. According to the characteristics of driver, a vehicle agent is modeled; the lane-changing rule and accelerated strategy of the vehicle agent are described; in the end, based on the agent technology, the multi-lane traffic flow is simulated and it indicated that the simulated traffic flow is coincident with the reality [18]. Considering the different vehicle destinations and types, a link layer controller for hierarchical automated highway system (AHS) architecture is presented, which is based on a principle of vehicle conservation and Lyapunov stability. The controller proposed stabilizes the vehicular destiny and flow around predetermined profiles in a stretch of highway using speed and lane changes as control signal. The simulation results indicate both, the validity of the assumptions about the dynamics of the coordination and regulation layers and the effectiveness of link lawyer controller [19].

On-ramp Metering Algorithm

Objective

Taking the lane difference into account, the error function $J(t)$ is defined as follows:

$$J(t) = \sum_{x=1}^{l} \lambda(x)|\rho(t,x) - \rho_c(x)| + \lambda_{\text{ramp}}\rho_{\text{ramp}}(t) \qquad (1)$$

where $x(x = 1, 2, \ldots l)$ is the xth lane. When $x = 1$, it represents the inner lane, while $x = l$ represents the outer lane. $\rho(t,x)$ is the density of the xth lane at time t. $\rho_c(x)$ is the expected density of the xth lane. $\rho_{\text{ramp}}(t)$ is the density of on-ramp at time t. $\lambda(x)$ is the regulator parameter of the xth lane. λ_{ramp} is the regulator parameter of ramp and $\sum \lambda(x) + \lambda_{\text{ramp}} = 1$.

To minimize the error function $J(t)$, a homogeneous linear differential equation of first order is formulated as follows, where $k > 0$. When error function $J(t)$ gets smaller, each density of mainstream lane will get close to the expected density while reducing the queue length at the same time.

$$kJ(t) + J'(t) = 0, \ J(t) = J(0)\,e^{-kt} \qquad (2)$$

To solve the error function $J(t)$, its derivative $J'(t)$ needs to be defined. And the traffic flow density $\rho(t, x)$ of each lane is required to be precisely detected. Thus, the derivative $J'(t)$ of the error function $J(t)$ can be written as:

$$J'(t) = \sum_{x=1}^{l} \lambda(x)[\rho(t, x) - \rho_c(x)] \cdot \rho'(t, x) + \lambda_{\text{ramp}} \rho'_{\text{ramp}}(t) \qquad (3)$$

where $\rho'(t, x)$ is the differential function of the xth lane density at time t. $\rho'_{\text{ramp}}(t)$ is the differential function of ramp density at time t.

To simplify the reasoning process, we choose the two-lane situation first. Define lane 1 as the inner lane and lane 2 as the outer lane:

$$J(t) = \lambda(1)|\rho(t, 1) - \rho_c(1)| + \lambda(2)|\rho(t, 2) - \rho_c(2)| + \lambda_{\text{ramp}} \rho_{\text{ramp}}(t) \qquad (4)$$

$$J'(t) = \lambda(1)[\rho(t, 1) - \rho_c(1)] \cdot \rho'(t, 1) + \lambda(2)[\rho(t, 2) - \rho_c(2)] \\ \cdot \rho'(t, 2) + \lambda_{\text{ramp}} \rho'_{\text{ramp}}(t) \qquad (5)$$

Algorithm

Fully considered the vehicle's safe distance, speed and route selection, and combined with the reality of vehicle lane-changing in multi-lane, TAN M put forward a dynamic discrete model of freeway traffic flow with multi-lane [16, 17]. Here, we define the lane density dynamic equation as follows:

$$\rho^{rs}(t+1, i, l) = \rho^{rs}(t, i, l) + \frac{T}{L_i}[q^{rs}(t, \varphi(i, l)) - q^{rs}(t, i, l) + u^{rs}(t, i, l) - e^{rs}(t, i, l)] \\ + \sum_{y \in C_i} \eta_{yl}^{rs}(t, i) \rho^{rs}(t, i, y) - \sum_{y \in C_i} \eta_{ly}^{rs}(t, i) \rho^{rs}(t, i, l)$$

$$(6)$$

where rs is the origin and destination of vehicle. (t, i, l) is the lth lane of road section i at time t. C_i is the set of lanes of road section i, and $l \in C_i$. L_i is the length of road section i. T is the cycle of data collection. $\rho^{rs}(t, i, l)$ is the average density of lth lane in origin-destination rs at time t. $u^{rs}(t, i, l)$ is the value that the vehicle number merging into (i, l) lane from on-ramp divides T at time segment $[(tT, (t+1)T)]$. $e^{rs}[(t, i, l)]$ is the value that the vehicle number leaving (i, l) lane and entrancing the off-ramp at segment $[(tT, (t+1)T)]$, and $q^{rs}(t, i, l)$ is the value that the vehicle number leaving (i, l) lane and flowing downstream divides T at segment $[tT, (t+1)T]$. $\eta_{yl}^{rs}(t, i)$ is the proportion of the vehicle changing lane from yth lane to lth lane in total vehicle number at road section i. $\varphi(i, l)$ is the lane set closed to (i, l) lane at upstream section $i - 1$.

This model considers vehicle's origin-destination, on-ramp, and off-ramp, in order to consider the particularity of vehicle changing-lane behavior, the model is simplified. We just design a two lanes in basic section, and a bottleneck section is not contained; we define inner lane as lane 1 and outer lane as lane 2, and then, the dynamic destiny equations of lane 1 and lane 2 are established, respectively, as follows:

$$
\begin{aligned}
\rho(t+1,i,1) &= \rho(t,i,1) + \frac{T}{L_i}[q(t,i-1,1) - q(t,i,1) + n_{21}(t,i) - n_{12}(t,i)] \\
\rho(t+1,i,2) &= \rho(t,i,2) + \frac{T}{L_i}[q(t,i-1,2) - q(t,i,2) + n_{12}(t,i) - n_{21}(t,i)]
\end{aligned}
\tag{7}
$$

Extending the two lanes to multi-lane (the lane number is constant) and concerning the vehicle's changing-lane behavior, each lane's dynamic destiny equation is written as:

$$
\begin{aligned}
\rho(t+1,i,l) &= \rho(t,i,l) + \frac{T}{L_i}\left[q(t,i-1,l) - q(t,i,l) + \sum_{y \in C_i} n_{yl}(t,i) - \sum_{y \in C_i} n_{ly}(t,i)\right] \\
&= \rho(t,i,l) + \frac{T}{L_i}[q(t,i-1,l) - q(t,i,l)] + \sum_{y \in C_i} \eta_{yl}(t,i)\rho(t,i,y) \\
&\quad - \sum_{y \in C_i} \eta_{ly}(t,i)\rho(t,i,l)
\end{aligned}
\tag{8}
$$

If one on-ramp is added to the simplified mode, each lane's dynamic destiny equation is:

$$
\begin{aligned}
\rho(t+1,i,l) &= \rho(t,i,l) + \frac{T}{L_i}\left[q(t,i-1,l) - q(t,i,l) + u(t,i,l) + \sum_{y \in C_i} n_{yl}(t,i) - \sum_{y \in C_i} n_{ly}(t,i)\right] \\
&= \rho(t,i,l) + \frac{T}{L_i}[q(t,i-1,l) - q(t,i,l) + u(t,i,l)] + \sum_{y \in C_i} \eta_{yl}(t,i)\rho(t,i,y) \\
&\quad - \sum_{y \in C_i} \eta_{ly}(t,i)\rho(t,i,l)
\end{aligned}
\tag{9}
$$

In practical traffic scenarios, the traffic densities are different in inner and outer lane; we design a traffic density measurement method based on input–output counting method; in this method, we select a road section just containing two lanes (the on-ramp is not considered temporarily) and consider the vehicle's changing-lane behavior. Detector is used to collect the vehicle data entrancing and exiting the area in unit time Δt; at same time, the real-time video supervision is used to obtain the vehicle's changing-lane behavior in unit time Δt, including the vehicle number of outer lane changing to inner lane and inner lane changing to outer lane. Fig. 1 shows traffic density measurement method based on input–output counting method.
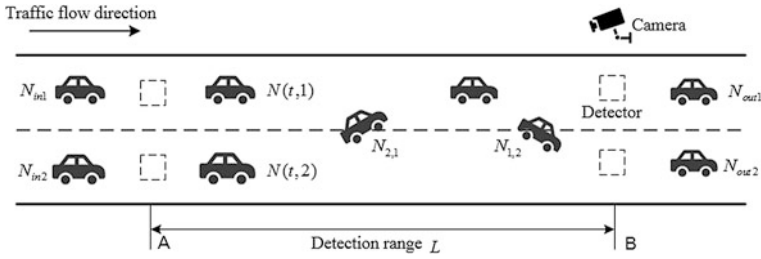
Fig. 1 Traffic density measurement method based on input–output counting method

So inner and outer lane's densities based on above measurement method are expressed as:

$$\rho(t, 1) = \frac{N(t, 1)}{L}, \ \rho(t, 2) = \frac{N(t, 2)}{L} \tag{10}$$

$$\rho(t + \Delta t, 1) = \frac{N_{\text{in1}} + N(t, 1) + N_{2,1} - N_{1,2} - N_{\text{out1}}}{L}$$
$$\rho(t + \Delta t, 2) = \frac{N_{\text{in2}} + N(t, 2) + N_{1,2} - N_{2,1} - N_{\text{out2}}}{L} \tag{11}$$

where $N(t, 1)$ is the initial vehicle numbers of inner lane at time t. $N(t, 2)$ is the initial vehicle numbers of outer lane at time t. L is the section length. $\rho(t + \Delta t, 1)$ is the section density of inner lane at time $t + \Delta t$. N_{in1} is the vehicle number entrancing inner lane in time Δt. $N_{2,1}$ is the vehicle number changing from outer lane to inner lane. $N_{1,2}$ is the vehicle number changing from inner lane to outer lane. N_{out1} is the vehicle number exiting outer lane in time Δt. $\rho(t + \Delta t, 2)$ is the section density of outer lane at time $t + \Delta t$. N_{in2} is the vehicle number of entrancing outer lane in time Δt. N_{out2} is the vehicle number exiting outer lane in time Δt.

In order to explore ramp metering algorithm in case of two lanes, we propose a simple on-ramp control model just including two lanes, and detectors are measured, respectively, in position A, B, ramp entrance and ramp stop line; meanwhile, the section length between mainline detectors is defined z, and section distance between on-ramp detectors is defined w. If it is assumed that on-ramp flow only merges into outer lane, but not inner lane, then the detection of outer lane's density needs to consider vehicle's changing-lane behavior and converging vehicle on ramp, but vehicle's changing-lane behavior is considered in inner lane. Based on the density measurement method, the traffic densities of inner and outer lane are collected, and video supervision is used to monitor the vehicle's changing-lane behavior of mainline (Fig. 2).

The vehicle of on-ramp can only merge to the outer lane of mainstream. Based on the inner and outer lane density formulas, the density of inner lane, outer lane, and on-ramp at time t is defined as follows:
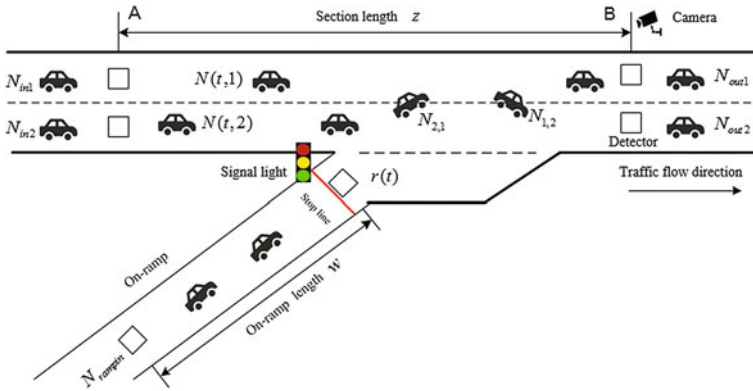
Fig. 2 On-ramp model contains two lanes of mainstream

$$\rho(t, 1) = \frac{N(t, 1)}{z}, \rho(t, 2) = \frac{N(t, 2)}{z}, \rho_{\mathrm{ramp}}(t) = \frac{N_{\mathrm{ramp}}(t)}{w} \qquad (12)$$

where $N_{\mathrm{ramp}}(t)$ is the initial vehicle number of on-ramp at time t.

The vehicle number of entering the inner lane and outer lane is N_{in1} and N_{in2}, respectively. N_{out1} and N_{out2} are the vehicle number of leaving the inner lane and outer lane. And $N_{1,2}$ is defined as the vehicle number of changing lane from the inner lane to the outer lane. On the contrary, $N_{2,1}$ is the vehicle number of changing lane from the outer lane to the inner lane. If the ramp metering rate at time t is $r(t)$, the vehicle number of merging the outer lane during the time interval Δt will be $r(t) \cdot \Delta t$.

$$
\begin{aligned}
\rho(t + \Delta t, 1) &= \rho(t, 1) + \frac{N_{\mathrm{in1}} + N_{2,1} - N_{1,2} - N_{\mathrm{out1}}}{z} \\
\rho(t + \Delta t, 2) &= \rho(t, 2) + \frac{N_{\mathrm{in2}} + N_{1,2} + r(t) \cdot \Delta t - N_{2,1} - N_{\mathrm{out2}}}{z}
\end{aligned}
\qquad (13)
$$

N_{rampin} is defined as the vehicle number of entering on-ramp during the time interval Δt. We obtain the following function of the on-ramp density at time $t + \Delta t$:

$$\rho_{\mathrm{ramp}}(t + \Delta t) = \rho_{\mathrm{ramp}}(t) + \frac{N_{\mathrm{rampin}} - r(t) \cdot \Delta t}{w} \qquad (14)$$

We have to deduce the differential functions $\rho'(t, 1)$, $\rho'(t, 2)$, and $\rho'_{\mathrm{ramp}}(t)$ to minimize the error function $J(t)$. And $\rho'(t, 1)$, $\rho'(t, 2)$, and $\rho'_{\mathrm{ramp}}(t)$ are defined as the differential function of the inner lane's, outer lane's, on-ramp's density function, respectively.

$$\rho'(t,1) = \frac{\rho(t+\Delta t,1) - \rho(t,1)}{\Delta t} = \frac{N_{\text{in1}} + N_{2,1} - N_{1,2} - N_{\text{out1}}}{z \cdot \Delta t} \tag{15}$$

$$\rho'(t,2) = \frac{\rho(t+\Delta t,2) - \rho(t,2)}{\Delta t} = \frac{N_{\text{in2}} + N_{1,2} - N_{2,1} - N_{\text{out2}}}{z \cdot \Delta t} + \frac{r(t)}{z} \tag{16}$$

$$\rho'_{\text{ramp}}(t) = \frac{\rho_{\text{ramp}}(t+\Delta t) - \rho_{\text{ramp}}(t)}{\Delta t} = \frac{N_{\text{rampin}}}{w \cdot \Delta t} - \frac{r(t)}{w} \tag{17}$$

Using Eqs. 5, 15, 16, and 17, we obtain the following derivative of error function $f'(t)$:

$$\begin{aligned}
J'(t) = {}& \lambda(1)[\rho(t,1) - \rho_c(1)] \cdot \frac{N_{\text{in1}} + N_{2,1} - N_{1,2} - N_{\text{out1}}}{z \cdot \Delta t} + \lambda(2)[\rho(t,2) - \rho_c(2)] \\
& \cdot \frac{N_{\text{in2}} + N_{1,2} - N_{2,1} - N_{\text{out2}}}{z \cdot \Delta t} + \lambda(2)[\rho(t,2) - \rho_c(2)] \cdot \frac{r(t)}{z} + \lambda_{\text{ramp}} \frac{N_{\text{rampin}}}{w \cdot \Delta t} - \lambda_{\text{ramp}} \frac{r(t)}{w}
\end{aligned} \tag{18}$$

So the ramp metering rate $r(t)$ of two lanes is:

$$\begin{aligned}
r(t) = {}& \frac{z \cdot w}{\lambda(2)[\rho(t,2) - \rho_c(2)] \cdot w - \lambda_3 \cdot z} \left\{ J'(t) - \frac{\lambda(1)[\rho(t,1) - \rho_c(1)] - \lambda(2)[\rho(t,2) - \rho_c(2)]}{z \cdot \Delta t} \right. \\
& \cdot (N_{2,1} - N_{1,2}) - \frac{\lambda(1)[\rho(t,1) - \rho_c(1)](N_{\text{in1}} - N_{\text{out1}}) + \lambda(2)[\rho(t,2) - \rho_c(2)](N_{\text{in2}} - N_{\text{out2}})}{z \cdot \Delta t} \\
& \left. - \frac{\lambda_{\text{ramp}} N_{\text{rampin}}}{w \cdot \Delta t} \right\}
\end{aligned} \tag{19}$$

The ramp metering rate equation of two lanes is extended to the case of multi-lane. And we define that vehicle can only change lane to the adjacent lane. In the case of multi-lane, the ramp metering rate equation for considering the mainline lane changing is:

$$\begin{aligned}
r(t) = {}& \frac{z \cdot w}{\lambda(l)[\rho(t,l) - \rho_c(l)] \cdot w - \lambda_{\text{ramp}} \cdot z} \cdot \left\{ J'(t) - \right. \\
& \frac{\sum_{x=1}^{l-1} \lambda(x)[\rho(t,x) - \rho_c(x)] - \lambda(x+1)[\rho(t,x+1) - \rho_c(x+1)]}{z \cdot \Delta t} \cdot \left(N_{x+1,x} - N_{x,x+1} \right) \\
& - \frac{\lambda(l)[\rho(t,l) - \rho_c(l)] - \lambda(1)[\rho(t,1) - \rho_c(1)]}{z \cdot \Delta t} \cdot \left(N_{1,l} - N_{l,1} \right) \\
& \left. - \frac{1}{z \cdot \Delta t} \sum_{x=1}^{l} \lambda(x)[\rho(t,x) - \rho_c(x)](N_{\text{inx}} - N_{\text{outx}}) - \frac{\lambda_{\text{ramp}} N_{\text{rampin}}}{w \cdot \Delta t} \right\}
\end{aligned} \tag{20}$$

where $N_{x+1,x}$ is the vehicle number of changing lane from the $x + 1$th lane to the xth lane. $N_{\text{in}x}$ is the vehicle number of entering the xth lane during the time interval Δt. $N_{\text{out}x}$ is the vehicle number of leaving the xth lane during the time interval Δt.

In order to simplify the equation, we define the function as $\partial(x) = \lambda(x)$ $[\rho(t, x) - \rho_c(x)]$. The lane-changing rate $\eta_{x+1,x}$ is defined as the ratio of the vehicle number that changes lane from the $x + 1$th lane to the xth lane and the total vehicle number of the $x + 1$th lane during the time interval Δt. So (20) can be written as:

$$r(t) = \frac{z \cdot w}{\partial(l) \cdot w - \lambda_{\text{ramp}} \cdot z} \left\{ J'(t) - \frac{\sum_{x=1}^{l-1} \partial(x) - \partial(x+1)}{\Delta t} \cdot \left[\eta_{x+1,x} \rho(t, x+1) - \eta_{x,x+1} \rho(t, x) \right] \right.$$
$$\left. - \frac{\partial(l) - \partial(1)}{\Delta t} \cdot \left[\eta_{1,l} \rho(t, 1) - \eta_{l,1} \rho(t, l) \right] - \frac{1}{z \cdot \Delta t} \sum_{x=1}^{l} \partial(x)(N_{\text{in}x} - N_{\text{out}x}) - \frac{\lambda_{\text{ramp}} N_{\text{rampin}}}{w \cdot \Delta t} \right\}$$
$$(21)$$

Simulation and Discussion

Simulation

In this section, we test the performance of the proposed MRD-RM algorithm using the micro-simulator VISSIM. The applicability and effectiveness of the MRD-RM algorithm are investigated and demonstrated. The simulation is conducted for a mono-directional, homogeneous two lanes of freeway mainstream with a single on-ramp. The simulation model is built based on the Fuqing on-ramp of Chengdu 2nd Ring Elevated Road. The interval distance z between mainstream detectors and the interval distance w between on-ramp detectors are set equal to 0.5 and 0.3 km, respectively. The time interval Δt used for all controllers in all investigations amounts to 30s. The simulation investigation has two comparison case based on ALINEA and no-control for providing a reference value. Simultaneously, VISSIM COM (Component Object Model) API (Application Programming Interface) is used to implement the detailed control scenarios of MRD-RM algorithm. And the implementation of ALINEA is based on VisVAP module. An optimal occupancy threshold 0.30 is calibrated to achieve the optimal performance of ALINEA.

The simulation model is established and calibrated based on field data collected during 5:00 PM to 7:00 PM at the Fuqing on-ramp segment of Chengdu 2nd Ring Elevated Road on September 10, 2015. The field data are then processed into traffic counts as inputs for VISSIM. We set the traffic flow of mainstream and on-ramp as 1890 and 607 veh/h, respectively. The desired density of line 1 and line 2 is $\rho_c(1)=36$ veh/km and $\rho_c(2)=37.5$ veh/km, respectively. Moreover, $\lambda(1) = 0.4$, $\lambda(2) = 0.45$, $\lambda(3) = 0.15$, and $k = 0.95$.
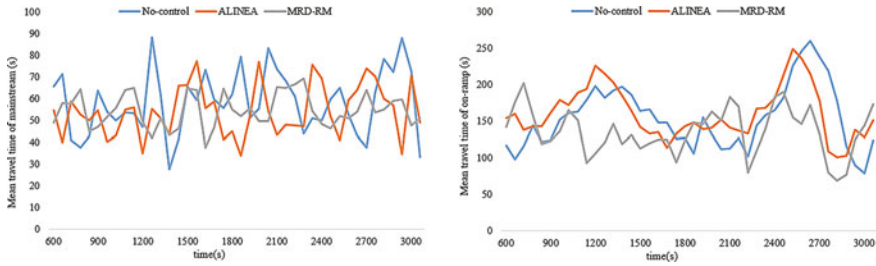
Fig. 3 Mean travel time of mainstream and on-ramp

Table 1 Mean output data of simulation

Index	No-control	ALINEA	Improvement rate (%)	MRD-RM	Improvement rate (%)
Mean speed of mainstream (km/h)	32.3	35.7	10.53	35.3	9.29
Mean travel time of mainstream (s)	57.5	52.6	−8.52	53.1	−7.65
Mean travel time of on-ramp (s)	153.9	159.5	3.64	131.5	−14.55
Mean vehicle delay of mainstream (s)	21.3	18.2	−14.55	18.4	−13.62
Mean vehicle delay of on-ramp (s)	83.1	98.6	18.65	75.5	−10.06
Mean queue length of on-ramp (m)	110.7	118.2	6.78	83.5	−24.57

We run the simulation with five random seeds, and the time length of each simulation investigation is 1 h. The results are evaluated by comparing mean speed of mainstream, mean travel time of mainstream and on-ramp, mean vehicle delay of mainstream and on-ramp, and average queue length of on-ramp. The time interval of counting simulation data is 60s during 600 to 3600s (Fig. 3).

Result Discussion

As we can observe from Table 1, both ALINEA and the proposed MRD-RM algorithm are capable of increasing mainstream speed and saving mainstream travel time. Seeing from the speed increase and time saving, ALINEA is slightly better than our MRD-RM algorithm. However, in the aspects of mean travel time, mean vehicle delay, and mean queue length of on-ramp, the MRD-RM algorithm shows significant improvement than ALINEA and non-control, while ALINEA shows no big difference with non-control.

Fig. 4 Density of
mainstream
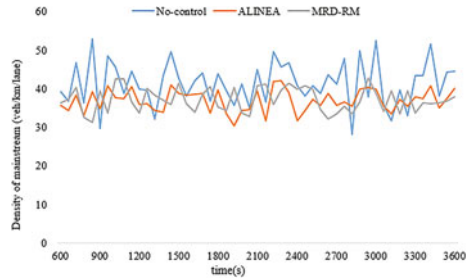
The main objective of the MRD-RM algorithm is keeping the lane density around the expected density. Thus, it can reach the similar achievement like ALINEA. Both MRD-RM and ALINEA can improve the mean speed of mainstream by 9.29 and 10.53%, respectively. The mean travel time of mainstream can be increased by 7.65 and 8.52%. The lower improvement of MRD-RM algorithm is mainly caused by synchronous consideration of mainstream and on-ramp queue. Thus, although the MRD-RM algorithm seems not to be the best control option compared with ALINEA, the gap of their optimal result is relatively small.

Nevertheless, seeing from the on-ramp sight, the effect of MRD-RM algorithm is significant. Mean travel time of on-ramp can be reduced by 14.55% as well as 10.06% drop of mean vehicle delay and 24.57% decrease of mean queue length of on-ramp. On the contrary, using ALINEA will increase 3.64% of the on-ramp mean travel time along with 18.65% of mean vehicle delay and 6.78% of mean queue length of on-ramp. At the same time, the curve of mainstream density will be smoother, which means the congestion condition is released when using MRD-RM, as shown in Fig. 4.

As one of the most important evaluation parameters, mean vehicle travel time can be reduced by using ALINEA and MRD-RM compared with non-control. Further, the performance of MRD-RM is the best since its optimal objective is minimizing on-ramp density.

Overall, the simulation result illustrates that MRD-RM algorithm can effectively reduce the queue length of on-ramp at the maximum limit while keeping the traffic volume of mainstream close to capacity. Compared with ALINEA and non-control, it surpasses them in the comprehensive optimal result showed both mainstream and on-ramp traffic control.

Conclusion

Based on the real-time density on-ramp control method, considering lane-change behavior among multi-lane mainstream in real situation of the expressway, this paper proposed a multi-lane real-time density-based on-ramp metering (MRD-RM) model. The minimization of density error function is presented as the optimal goal.

And the dynamic discrete model of multi-lane traffic flow character is introduced when calculating the expressway density. The micro-simulator VISSIM is used for testing the performance of the MRD-RM model setting ALINEA and non-control as the reference method. The simulation result shows our proposed MRD-RM algorithm is capable of reducing the queue length of on-ramp at the maximum limit while keeping the traffic volume of mainstream close to capacity.

On the other hand, when traffic demand of both mainstream and on-ramp is large, especially in the exceeding capacity situation, there is a trade-off between reducing ramp queue and maximizing mainstream traffic flow. How to balance these two sides and reaching the system optimization needs further discussion.

References

1. Papageorgiou, M., and Kotsialos, A. 2000. Freeway ramp metering: An overview. In *Proceedings of 3th IEEE Intelligent Transportation Systems Conference*, Dearborn Mich, 271–281.
2. Papageorgiou, M., E. Kosmatopoulos, Papamichail, I. et al. 2007. ALINEA maximises motorway throughput-an answer to flawed criticism. *Traffic Engineering and Control* 48 (6): 271.
3. Wan-bao, G., W. Jian, and Z. Jiao. 2012. On-ramp metering algorithm based on real-time density. *Journal of Transportation Systems Engineering and Information Technology* 12 (2): 150–155.
4. Scariza, J. R. 2003. Evaluation of coordinated and local ramp metering algorithms using microscopic traffic simulation. Massachusetts Institute of Technology.
5. Zhang Hai-jun, Yang Xiao-guang, and Zhuang Jue. 2006. Review of freeway on ramp metering methodologies. *Journal of Tongji University (Natural Science)* 33 (8): 1051–1055.
6. Masher, D. P., Ross, D. W., Wong P. J. et al. 1975. Guidelines for design and operation of ramp control systems.
7. Papageorgiou, M., H. Hadj-Salem, and Blosseville, J. M. 1991. ALINEA: A local feedback control law for on-ramp metering. Transportation Research Record, 1991 (1320).
8. Papageorgiou, M., H. Hadj-Salem, and Middelham, F. 1997. ALINEA local ramp metering: Summary of field results. *Transportation Research Record* 1603: 90–98.
9. Papageorgiou, M., J. M. Blosseville, and Haj-Salem, H. 1991. Modelling and real-time control of traffic flow on the southern part of Boulevard Périphérique in Paris: Part II: Coordinated on-ramp metering. *Transportation Research Part A: General* 24 (5): 361–370.
10. Smaragdis, E., and M. Papageorgiou. 2003. Series of new local ramp metering strategies. *Transportation Research Record* 1856: 74–86.
11. Smaragdis, E., M. Papageorgiou, and Kosmatopoulos, E. 2004. A flow-maximizing adaptive local ramp metering strategy. *Transportation Research Part B: Methodological* 38 (3): 251–270.

12. Wang, Y., Papageorgiou, M., Gaffney, J. et al. 2010. Local ramp metering in the presence of random-location bottlenecks downstream of a metered on-ramp. In *Intelligent Transportation Systems, 13th International IEEE Conference*, 1462–1467.
13. Yang, X., L. Chu, and Recker, W. 2002. GA-based parameter optimization for the ALINEA ramp metering control. In *Intelligent Transportation Systems, 5th IEEE International Conference*, 627–632.
14. Da-yan, W., T. Hui-li, K. Ling-jiang, and Mu-ren, L. 2005. Study on a three-lane cellular automata traffic flow model. *Journal of System Engineering* 20 (4): 393–397.
15. Hui-li, T., K. Ling-jiang, and Mu-ren. L. 2005. A cellular automation model for mixing traffic in two-lane system. *Acta Physica Sinica* 54 (8): 3516–3521.
16. Tan, M. 2003. Hierarchical optimal control problem and its algorithm based on traffic flow discrete model in multilane freeway. *Control Theory & Applications* 6: 855–858.
17. Tan, M., S. Tang, and Jian, X. 2002. Dynamic discrete traffic model of freeway with multiple lanes. *China Journal of Highway & Transport* 15 (2): 91–94.
18. Wei, L., H. Dong-zhi, and Run-mei, L. 2005. Distributed simulation of multi-lane traffic flow based on multi-agent. *Computer Simulation* 22 (2): 191–194.
19. Alvarez, L., R. Horowitz, and Yoy, C. V. 2003. Multi-destination traffic flow control in automated highway system. *Transportation Research Part C* 11: 1–28.

Vehicle Driving Characteristics on Rural Highways and the Evaluation of Stability Performance Based on Lorenz Scatter Plot

Xue-lian Zheng, Xian-sheng Li, Yuan-yuan Ren and Xiang-yu Meng

Abstract Rural highways with low technical standards take up a great proportion of the total highways in China, and the problem of traffic safety is quite server on it. To investigate vehicle driving stability on rural highways, vehicle driving state and trajectory were collected by field tests. It was discovered that there are six different kinds of driving trajectories when passing through curved sections on rural highways, and they are named as the ideal trajectory, the normal trajectory, the cutting trajectory, the swing trajectory, the drifting trajectory, and the correcting trajectory. To evaluate vehicle handling stability while drives along different trajectories, Lorenz scatter plot was used to discover the parameter that could reflect driving stability best of all. Then, vehicle driving stability on different trajectories was studied by CarSIM and it was found that vehicle has the best driving stability when drives along the cutting trajectory and has the worst stability while drives along the correcting one.

Keywords Rural highways · Vehicle driving stability · Performance evaluation · Lorenz scatter plot · Lateral acceleration

Introduction

Rural highways composed a great proportion of the total highways and played an important role in boosting economic development of rural areas. However, due to low technical grade and limited construction funds, rural highways usually have a poor road condition and alignment. In 2012, traffic accidents happened on rural

X. Zheng · X. Li · Y. Ren (✉) · X. Meng
School of Transportation, Jilin University, 130022 Changchun, China
e-mail: renyy@jlu.edu.cn

X. Zheng
e-mail: zhengxuelian@jlu.edu.cn

X. Li
e-mail: lixs@jlu.edu.cn

highways took up about 50% of the total traffic accidents in China, and had caused serious human injured and death [1]. Many researches had been done to investigate vehicle handling stability on roads with low technical standards, and it was discovered that horizontal curve is a significant safety issue. Statistical showed that the accident rate and severity in moderate curves are at their highest levels, which are 55 and 58%, respectively [2].

In 1960, Thomas measured vehicle driving stability on curved sections and found out road edge-line has great influence on vehicle driving trajectory. While turning left, vehicle driving trajectory trends to get close to road centerline and it trends to get away from road centerline while turning right [3]. In 1970, Gelnnon and Weaver done plenty of observation on vehicle driving trajectory and discovered that road curvature plays an important role on vehicle driving trajectory [4]. In 1997, Weise et al. studied the lateral deviation from vehicle driving trajectory to road centerline, and some important conclusions were drawn [5]. In 1999, Chatziastros et al. investigated the importance of velocity information during a lane-centering task in a simulated tunnel. It was found that drivers used both velocity and spatial frequency information to maintain a centered position on a path, and the presence of lane borders improved the accuracy of lane-centering [6]. In 2005, Professor Xiao-duan Sun done a research on lateral position of vehicle driving trajectory on seven different roads in America. Based on the observation results, it was discovered that lane width, road edge-line, and road condition all have an influence on vehicle lateral position. Besides, road edge-line could help hinting road alignment to driver and keeping vehicle drives along the road centerline [7]. Other researchers, such as Mester [8], Spacek [9–12], and so forth [13, 14], studied the distribution of vehicle position during curve driving and analyzed factors that could have an influence on vehicle driving trajectory. Some salient conclusions were drawn to help road construction.

Besides those researches, some researchers studied the relation between vehicle driving trajectory and road alignment from the view of driver's environment perception. Donges in 1978 [15], Beall in 1996 [16], Kageyama in 1991 [17], and so forth investigated the influence of road alignment and curvature on the distribution of drivers' gaze point, as well as the influence of the distribution of gaze point on vehicle driving trajectory and stability.

Therefore, the paper would do a research on vehicle driving stability on rural roads with low technical standards. Vehicle driving conditions and trajectory were collected by field tests. Types of driving trajectory on rural highways were analyzed. To evaluate vehicle handling stability while drives along different trajectories, Lorenz scatter plot was used to discover the parameter that could reflect vehicle driving stability. Then, by vehicle dynamic analysis, it was found that vehicle will have the best driving performance when drives along the cutting trajectory and the worst performance while drives along the correcting one.

Method and Material

Data Collection

To reflect the characteristics of rural highways, two typical rural highways in Jilin province were selected and used as experiment routine, as shown in Fig. 1.

The rural highway shown in Fig. 1 (1) has a total length of 15 km. U-turns, steep slopes, and combined bending slopes are all also included. It passes through several villages and is a standard 2-lane dual carriageway. Road width is 7.5 m and lane width is 3.5 m. On the curved sections, road width drops to 6 m. The design speed for this road is 40 km/h, and the speed limitation on turnings decreased to 20 km/h. Most of road markings assembled along this highway are blur, and some road parts even do not have any markings. Road signs are incomplete, and some U-turns do not have any warning signs on the condition of limited visual field. Traffic density is low on this road.

For the second one shown in Fig. 1 (2), it has a total length of 20 km. This highway connects town and distant villages, traffic density is much lower than the first one, and only a few villages locate along this road. The road width is 7.5 m and lane width is 3.5 m; some curved section has a width of 9 m. Road design speed is 40 km/h and it drops to 20 km/h on some U-turns. Road markings are not clear, and some road parts do not have any markings.

Two drivers were selected to drive a Mazda on the experiment routine for several times. Vehicle driving conditions were collected by VBOX and inertial sensor, including driving speed and trajectory, steering angle, brake force, slip angle, roll angle, roll rate, yaw rate, etc. No specific rules were obeyed by drivers.

Types of Vehicle Driving Trajectory on Rural Highways

Vehicle driving trajectory could reflect the balance of driver operation and vehicle dynamic response. Therefore, trajectory characteristics were analyzed for different
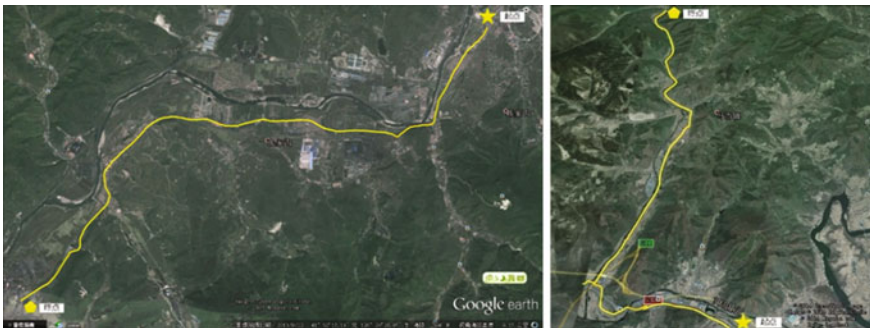
Fig. 1 Typical rural highways

drivers. According to the time when vehicle drives into or out a curved section, and the position where vehicle drives on the opposite lane, driving trajectory for different drivers could be divided into six types: the ideal trajectory, the normal trajectory, the cutting trajectory, the swing trajectory, the drifting trajectory, and the correcting trajectory. Details for the six types of driving trajectory and the corresponding mathematical models could be found in Ref. [13, 14].

Among the six kinds of vehicle driving trajectories, the radius of correcting trajectory is smaller than that of road centerline, and the radius is getting smaller with the radial derivation and the circumferential length increase. For the other kinds of driving trajectories, the radius of driving trajectory is larger than that of road centerline, and with the increase of characteristic parameters of driving trajectory, the radius gets larger.

The Evaluation Analysis of Vehicle Driving Stability on Different Trajectories

The rural highways have a much worse conditions than highways with high technical grades. Furthermore, rural highway passes through villages, and it brings a diverse types of road environment and traffic conditions. Therefore, vehicle driving velocity is much lower on rural highways, and vehicle's inherent technical condition has a little influence on its driving stability.

It had been studied for decades that driver plays a quite salient role in keeping vehicle handling stability. In the driver–vehicle–road closed-loop system, driver is the receiver of external information and the instruction giver of vehicle operation. He is the key for safety driving. By observing road environment and traffic conditions, combined with the command of vehicle driving performance, driver gives appropriate operations to the car. At the same time, continuous operation correcting is made to keep vehicle driving safely and fluently. There are many parameters that could reflect vehicle handling stability, such as yaw rate, roll angle and roll rate, slip angle, lateral velocity and acceleration, and so on. For the traditional evaluation of vehicle handling stability, a single driving parameter was picked out based on simplified vehicle dynamic analysis to describe vehicle handling stability in the open-loop system. Moreover, the simplified vehicle dynamic models usually express the liner part of vehicle dynamic characteristics. However, the closed-loop driver–vehicle–road system reflects a strong nonlinear characteristics and the traditional evaluation method of vehicle handling stability is not suitable anymore. A new method was proposed in this paper to discover parameters that could describe vehicle handling stability when driving on rural roads.

The Lorenz Scatter Plot

The Lorenz scatter plot was popularly used in the analysis of heart rate variability (HRV), and it does not require the HRV signal to be stationary [18, 19]. The Lorenz scatter plot is a diagram in which each RR interval is plotted against the previous one, as shown in Fig. 2. The figure was cited from Ref. [19].

The scatter in Fig. 2 could be expressed as follows:

$$\begin{cases} x_n = R(n) - R(n-1) \\ y_n = R(n+1) - R(n) \end{cases}, \tag{1}$$

where $R(n)$ is the time when the nth R peak happened in a continuous electrocardio signal. The time interval between two adjacent R peaks means the transient heartbeat cycle.

It was quite clear that point locates on the $x = y$ line means there is no difference between the current and previous RR intervals, and the point locates along the $x = -y$ line means there is a large difference on the current and previous RR intervals. The vertical distance from point to $x = y$ line could reflect the difference between current and previous RR intervals and also reflect the HRV degree. The larger the vertical distance is, the larger the HRV is.

As shown in Fig. 2, SD_1, the vertical deviation parameter, reflects the short-term variability. It was used to express the activity of the parasympathetic system. The parasympathetic system plays a quite important role when individual is relax. When

Fig. 2 Lorenz scatter plot for HRV
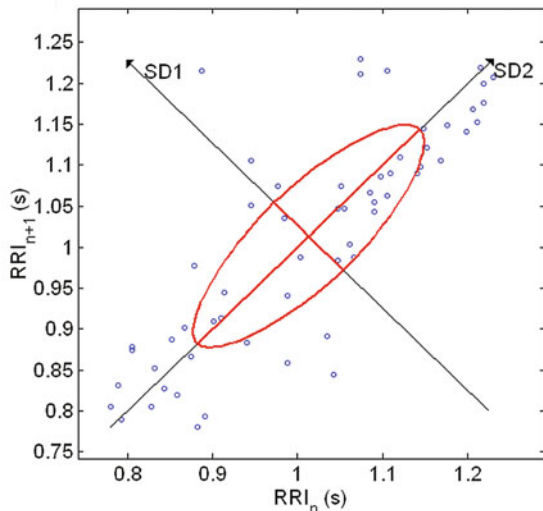
the SD_1 decrease, it means the parasympathetic system's activity decrease. SD_2, the horizontal deviation, reflects the long-term variability. It was used to express the activity of sympathetic system. The sympathetic system plays an important role when individual is excited or nervous. When the SD_2 decreases, it means the activity of sympathetic system increases. For a normal individual, there is an internal balance between the activity of sympathetic system and that of parasympathetic system, and the SD_1/SD_2 ratio will be changed in a certain range.

Generally, a Lorenz scatter plot for a normal individual will present as a comet shape which has high density and small vertical deviation on the lower left corner and has low density and relatively large vertical deviation on the top right corner.

The Lorenz Scatter Plot in Analysis of Vehicle Handling Stability

A typical time-varying curve for any one of vehicle driving parameters is presented in Fig. 3. Due to vehicle dynamic characteristics and the continuous operation correcting from drivers, the whole vehicle driving performance could be divided into two parts: the low-frequency part and the high-frequency part. The low-frequency part reflects the parameter's entire changing trend, such as increases or decreases progressively, as well as changes periodically. The high-frequency part reflects the parameter's unstable fluctuation, such as transient overshoot, vibration, and step response. Mapping to vehicle dynamic features, the low frequency will present the linear ingredient and the high frequency will present the nonlinear ingredient. The linear and nonlinear ingredients interact with each other, and vehicle handling stability is a comprehensive stable condition on both high-frequency and low-frequency parts.

Fig. 3 Time-varying curves for lateral acceleration
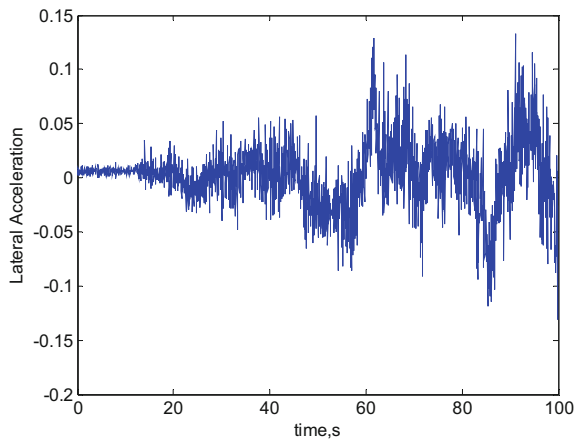
While the high-frequency parts fluctuate with a small amplitude and last for a quite short time, vehicle is transient unstable, and this transient unstable may lead to vehicle out of control. While the value of low-frequency parts grows progressively, vehicle will lose its stability at a certain time. However, the linear parts play a much more important role in deciding vehicle handling stability and it should be paid much more attention.

As a reference of Lorenz scatter plot used to analyze individual HRV, the method was cited to pick out the driving parameter that could describe vehicle driving stability vividly and practically. Therefore, a Lorenz scatter plot for an arbitrary vehicle driving parameter will be a diagram, in which the current change rate will be plotted against the previous one. The current and previous change rates for an arbitrary vehicle handling stability parameter could be expressed as

$$\begin{cases} x_n = f(n) - f(n-1) \\ y_n = f(n+1) - f(n) \end{cases}, \tag{2}$$

where f is an arbitrary vehicle driving parameters, and $f(n)$ is the nth value in a parameter's time series.

According to Eq. (2), change rate for vehicle driving parameters in a period of time could be obtained and so did the Lorenz scatter plot, as shown in Fig. 4. In Fig. 4, the rotation angle from x-axis to x'-axis is 45 degree, so does the rotation angle from y-axis to y'-axis.

As shown in Fig. 4, points along the x'-axis mean that the current change rate is quite similar to the previous one, and the parameter changes with constant acceleration or deceleration. The distribution range of points along the x'-axis reflects the low-frequency part of vehicle driving performance and the parameter's variation

Fig. 4 The Lorenz scatter plot
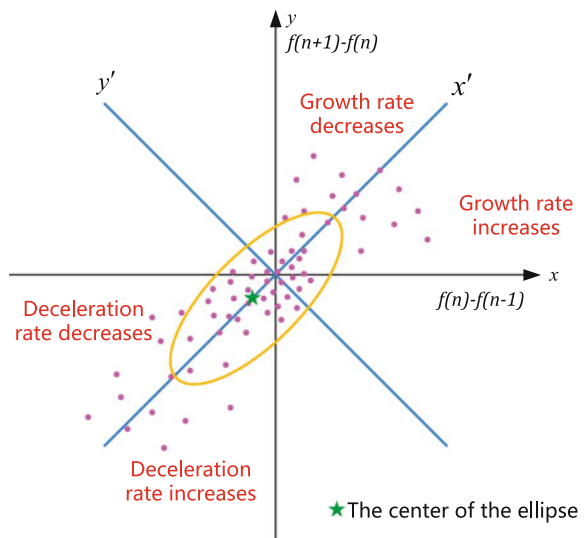
range during vehicle driving process. Points along the y'-axis means that the current change rate is not equal to the previous one, and the parameter value changes with no regular pattern. The distribution range of points along the y'-axis reflects the high-frequency part of vehicle driving performance and the rapid change composition of the parameter.

While the points locate in the first quadrant, it means the parameter value increases gradually. When the points locate in the third quadrant, it means the parameter value decreases gradually. If the points fall in the neighborhood of the origin, it means change rate of this parameter is equal to zero approximately.

It was quite obvious that the horizontal distribution range along the x'-axis expresses the smallest and largest change rates of parameters. A large horizontal distribution range shows that during vehicle driving process, the parameter changes violently. Besides, the more violent the parameter changes, the more vivid and effective the parameter reflects vehicle handling stability. Compared with the horizontal distribution range, the vertical distribution range along the y'-axis has a much less function to describe vehicle handling stability, but it does reflect the transient overshoot and fluctuations. Besides, the smaller vertical distribution and larger gathering density along the y'-axis indicate that vehicle drives more fluent and stable.

For expression simplicity, points in the x–y coordinate system should be transferred into the x'–y' coordinate system. The transformation could be expressed as

$$\begin{cases} x_2 = y_1 \sin \theta + x_1 \cos \theta \\ y_2 = y_1 \cos \theta - x_1 \sin \theta \end{cases}, \tag{3}$$

where (x_1, y_1) is point's coordination in the x–y coordinate system, (x_2, y_2) is the same point's coordination in the x'–y' coordinate system.

The centroid of points distribution in the x'–y' axis could be expressed as

$$\begin{cases} x_C = \dfrac{1}{m} \displaystyle\sum_{n=1}^{m} x_2(n) \\ y_C = \dfrac{1}{m} \displaystyle\sum_{n=1}^{m} y_2(n) \end{cases}, \tag{4}$$

where m is the number of points.

The horizontal deviation along the x'-axis is given as

$$SD_1 = \sqrt{\frac{1}{m} \sum_{n=1}^{n=m} (x_2(n) - x_C)^2} \tag{5}$$

and the vertical deviation along the y'-axis is given as

$$SD_2 = \sqrt{\frac{1}{m}\sum_{n=1}^{n=m}(y_2(n)-y_C)^2}. \tag{6}$$

An ellipse was painted based on the distribution of points, as shown in Fig. 4. Its center is the centroid of points distribution, the length of half major axis is equal to SD_1, and the length of half minor axis is equal to SD_2. An indicator was defined to describe the balance of high-frequency part and low-frequency part of vehicle dynamic analysis, which is presented as

$$ind = SD_2/SD_1. \tag{7}$$

For a stable driving vehicle and any one of vehicle driving parameters, the *ind* should be bigger than or equal to 1. Furthermore, parameter that has a larger *ind* would be more appropriate to reflect vehicle handling stability. Therefore, due to the nonlinear characteristics when vehicle drives on rural roads, the Lorenz scatter plot was used to discover parameter that could express vehicle handling stability.

Indicator to Reflect Vehicle Driving Stability

Vehicle driving parameters could be divided into two parts: the import ones and the export ones. Steering angle, brake force, and the depth of acceleration pedal are included in the import parameters, and these three parameters decide vehicle driving direction and speed. Lateral velocity and acceleration, yaw rate, slip angle, roll angle, and roll rate, as well as pitch angle and pitch rate, are included in the export parameters, and they reflect vehicle dynamic characteristics and its driving stability.

To investigate the parameter that could reflect vehicle driving stability when vehicle drives on rural highways, Lorenz scatter plots for lateral acceleration, yaw rate, roll angle, and roll rate are shown in Fig. 5.

(1) Roll angle　　　　(2) Roll rate　　　　(3) Yaw rate　　　　(4) lateral acceleration
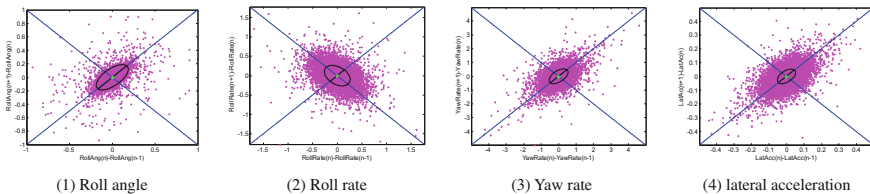
Fig. 5 Lorenz scatter plot for different parameters

It can be seen from Fig. 5 that among the four parameters, the Lorenz scatter plot of lateral acceleration has a much bigger *ind* value than the others. Besides, the distribution density of lateral acceleration is much bigger than the others. For roll rate, both the long-term and short-term deviations are quite small, and the distribution density is quite low.

In generally, roll angle, yaw rate, and lateral acceleration are the most important DOFs in establishing vehicle dynamic models. Therefore, roll angle, yaw rate, and lateral acceleration are the most widely used parameters to reflect vehicle driving stability. Hence, *ind* values for roll angle, roll rate, lateral acceleration, and yaw rate when vehicle drives on straight and curved section with different radii were calculated. The calculated results are presented in Table 1.

It can been drawn from Table 1 that when vehicle drives on straight or curved sections with much smaller curvatures, *ind* values for roll angle, roll rate, yaw rate, and lateral acceleration are all quite small. The values are approximately equal to 1, which means that the long-term and short-term deviations for vehicle driving parameters are quite similar to each other. That is to say, in the condition of vehicle drives on rural highways with quite low driving speed and relatively small steering angle, driver needs to do a lot of amendments to keep vehicle drives safely. However, due to low driving speed, vehicle driving stability is quite good. With the increase of road curvature, vehicle driving stability drops gradually. When vehicle drives on curved sections whose radius is in the range of 100–160 m, roll rate has the largest *ind* value. While vehicle drives on curved sections whose radius is smaller than 100 m, *ind* values for four parameters are all quite large, especially for roll angle, yaw rate, and lateral acceleration. Among the selected four parameters,

Table 1 *ind* values for roll angle, roll rate, yaw rate, and lateral acceleration on kinds of road sections

Road section	Turning radius (m)	Roll angle	Yaw rate	Roll rate	Lateral acceleration
Straight road	–	1.776	*2.569*	1.409	1.196
Turn right	Above 600	0.999	1.167	*1.657*	1.013
	350–600	1.001	1.069	*1.085*	1.003
	220–350	0.921	*1.825*	1.412	0.915
	160–220	1.011	*7.562*	1.082	0.957
	100–160	1.040	1.090	*2.979*	0.999
	60–100	24.584	38.655	2.776	*55.579*
Turn left	Above 600	0.999	1.032	*1.121*	1.003
	350–600	0.996	0.985	*1.105*	1.000
	220–350	0.983	*1.717*	1.156	0.951
	160–220	0.989	*1.260*	1.062	0.962
	100–160	0.991	1.472	*2.698*	0.972
	60–100	23.286	23.353	3.014	*71.393*

lateral acceleration has the biggest *ind* value, which means that lateral acceleration has the largest long-term deviation, and it is the most import part that driver should be pay attention on to keep vehicle driving stability.

For a more clear observation, *ind* values for lateral acceleration, roll rate, yaw rate, and roll angle when vehicle drives on straight and on curved sections are shown in Fig. 6.

It can be seen from Fig. 6 that when vehicle drives on straight or curved sections whose radius is quite large, the evolution trend of vehicle driving performance is quite mild, vehicle's dynamic respond to certain input is relatively gentle, and the long-term deviation of vehicle driving stability is limited in a narrow range. Therefore, no conclusion could be drawn on which parameter could be used to judge vehicle handling stability However, when vehicle drives on curved sections whose curvature is quite big, *ind* values for yaw rate, roll angle, and lateral acceleration are much bigger, and *ind* value for lateral acceleration even grows to more than 50. That is to say, vehicle's lateral acceleration changes in a great range while vehicle passes through U-turns and the change of lateral acceleration could reflect vehicle handling stability. Except for lateral acceleration, yaw rate and roll angle are also changed in a wide range when vehicle drives along curved section with small radius. Therefore, lateral acceleration, roll angle, and yaw rate are the parameters that used to distinguish vehicle driving stability and to evaluate vehicle handling stability when drives on rural roads.

Vehicle Driving Stability on Different Types of Driving Trajectories

To investigate vehicle driving stability when drives on six different kinds of driving trajectories, the trajectories were established in the CarSim software. The settings of road superelevation, cross slope, lane width, and attachment coefficient were based on road construction standards. Vehicle driving speed was equal to road design speed, which is 40 km/h.

Time-varying curves for lateral acceleration, roll angle, yaw rate, and roll rate are shown in Fig. 7.

It can be seen from Fig. 7 that vehicle driving stability when drives on different kinds of trajectories differs greatly from each other. While driving along road centerline, vehicle has the worst driving stability; and while driving along the cutting trajectory, vehicle has the best driving stability. While vehicle drives on the other trajectories, the order of vehicle driving stability is swing trajectory, drifting trajectory, normal trajectory, and ideal trajectory. Besides, vehicle driving stabilities on swing and drifting trajectories are quite similar.

Trajectory lengths for six different kinds are listed in Table 2.
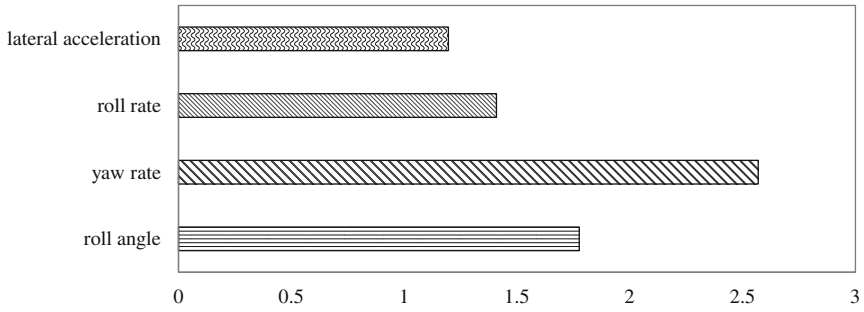
(1) *ind* values on straight sectinos
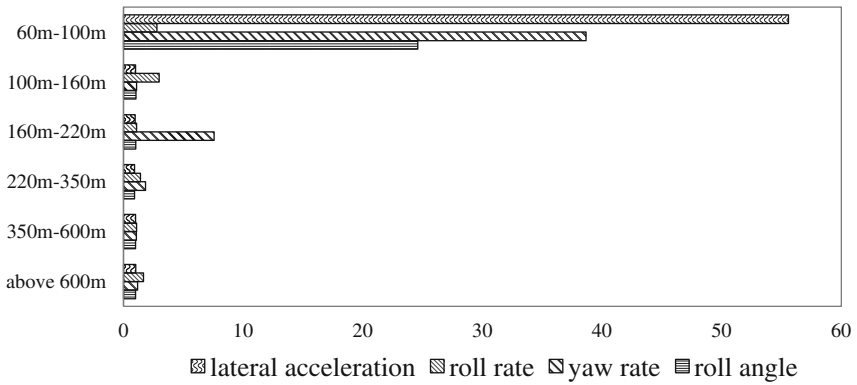
(2) *ind* values on right turning sections
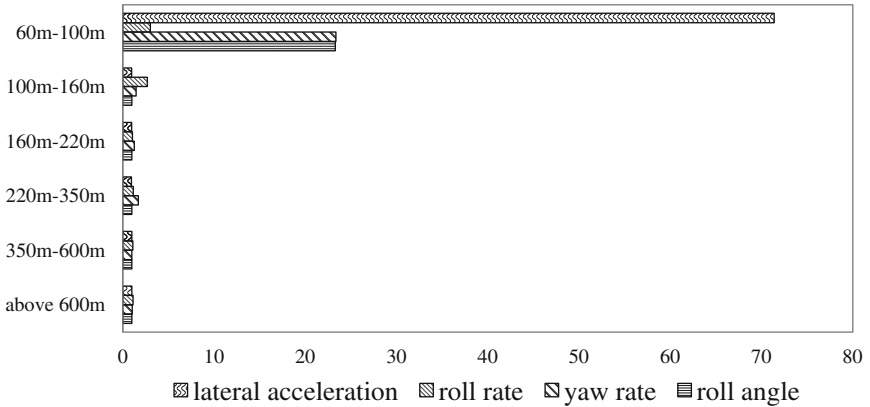
(3) *ind* values on left turning sections

Fig. 6 *ind* values when vehicle drives on different road sections

(1) roll rate   (2) roll angle   (3) yaw rate   (4) lateral acceleration
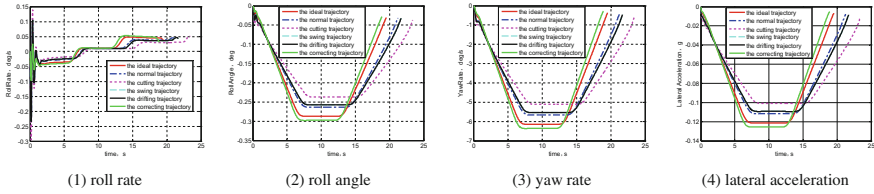
Fig. 7 Vehicle driving performance on six trajectories

Table 2 Length for six different kinds of driving trajectories

Trajectories	The ideal one	The normal one	The cutting one	The swing one	The drifting one	The correcting one
Length, m	217	236	260	241	241	210

It was drawn from Table 2 that vehicle driving stability on certain trajectory is proportional to its length. Therefore, under the situation of same arc angle, trajectory with large radius will ensure vehicle having good stability performance.

Discussion

While drives on rural highways, vehicle driving speed is much lower, and its driving stability could not be evaluated by driving speed. Under the condition of limited horizon, driver could not have a good command of road alignment and curvature. Therefore, during the curved section driving, driver will correct their operation from time to time to ensure vehicle could pass through bends safely and quickly. Based on different driving experiences, environment perceptions, and vehicle control abilities, vehicle driving trajectories are quite different from person to person.

By field tests, vehicle driving trajectory when passes through curved sections was collected and analyzed. According to the time when vehicle entering and existing curved sections, and the position when vehicle drives on the opposite lane to increase the radius of driving trajectory, vehicle driving trajectory for different drivers could be divided into six types: the ideal trajectory, the normal trajectory, the cutting trajectory, the swing trajectory, the drifting trajectory, and the correcting trajectory. Due to the lack of road signals or driving instructions, driver could not have an comprehensive command on road alignment and could not judge the start or ending point of curved section, and therefore, sometimes vehicle already enters a bend until driver realized that. To ensure driving safety, driver needs to correct their operation to vehicle from time to time. Furthermore, by taking up the opposite lane,

driver wants to have a good vision field to command road alignment, and to increase the radius of driving trajectory.

Among the six different driving trajectories, the normal, the ideal, and the correcting one do not intersect with road centerline, which means that vehicle drives on its own lane and does not take up on the opposite lane during curves driving. The swing and drifting trajectories intersect with road centerline once, and vehicle drives on the opposite lane in the entering or existing part of curved section. The cutting trajectory intersects with road centerline twice, and vehicle drives on the opposite lane in middle part of curved section.

It was known from vehicle dynamic analysis that the cutting trajectory has the biggest length, and vehicle driving stability is the best when vehicle drives along the cutting trajectory. This result shows that by driving on the opposite lane in middle part of the curved section, vehicle handling stability could be greatly improved. It was also drawn from the simulation results that the correcting trajectory has the smallest length, and vehicle driving stability is the worst when vehicle drives along the correcting trajectory. The result shows that inappropriate amendment and decision lead to poor vehicle driving performance. Hence, driving experience, environment perception, and decision-making method will have an important influence on vehicle handling stability. To ensure vehicle driving safety, road signs and signals, as well as necessary driving instructions, are quite important.

Conclusions

Different from the traditional vehicle handling stability analysis, vehicle driving stability on rural roads presents a strong nonlinear characteristic. Hence, to find out indicator that could evaluate vehicle handling stability, the Lorenz scatter plot that has a quite popular application in HRV analysis and could reflect system's nonlinear features was used. The ratio of long-term and short-term deviations of parameters was defined to describe vehicle driving stability. The bigger the ratio is, the much more effective the parameter describes vehicle driving stability. According to analysis results, lateral acceleration is the most effective one to describe vehicle driving stability when drives on curved sections, and yaw rate and roll angle come next.

The six different trajectories were established in CarSim software, and dynamic simulation was carried out. According to simulation results, the cutting trajectory has the greatest vehicle driving stability and the correcting one has the worst vehicle driving stability. Besides, the cutting trajectory has the largest length and the correcting trajectory has the smallest one. It was concluded from trajectory characteristics that driving experience, environment perception, and decision-making method will have an important influence on vehicle handling stability. To ensure vehicle driving safety, road signs and signals, as well as necessary driving instructions, are quite important. Therefore, how to improve vehicle driving stability by the setting of road signs and signals will be the study in the near future.

Conflict of interests The authors declare that there is no conflict of interests regarding the publication of this article.

References

1. Research Institute of Highway, Ministry of Transport, P.R. China. 2013. *The blue book of road safety in China*. Beijing: China Communications Press.
2. Charlton, S.G., and J.J.D. Pont. 2007. *Curved speed management*. Research Report 323, Land Transport New Zealand, Wellington, New Zealand.
3. Thomas, I.L., and W.T. Taylor. 1960. Effect of edge striping on traffic operations. *Highway Research Board Bulletin* 244:11–15.
4. Glennon, J.C., and G.D. Weaver. 1971. *The relationship of vehicle paths to highway curve design*. Research Report 134–5, Texas Transportation Institute, Texas.
5. Weise, G.S., R. Teter, A. Sossumithen, et al. 1997. *Lateral placement of vehicles on curves of two-lane rural roads as safety criterion*. Toronto, Canada: XIIIth World meeting of the International Road Federation.
6. Chatziastros, G., Wallis, M., and H.H. Bülthoff. 1999. *The use of optical flow and splay angle in steering a central path*. Technical Report No. 72, Max Planck Institute, Germany.
7. Sun, X.D., and T. Dean. 2005. *Impact of edge lines on safety of rural two-lane highways*. Technical Report 414, Civil Engineering Department University of Louisiana at Lafayette, Lafayette, La, USA, Oct 2005.
8. Mestre, D.R. 2001. Dynamic evaluation of the useful field of view in driving. In *Proceedings of the first international driving symposium on human factors in driving assessment*, *training and vehicle design*, 234–239.
9. Spacek, P. 2005. Track behavior in curve areas: Attempt at typology. *Journal of Transportation Engineering* 131 (9): 669–676.
10. Scheifele, U., and P. Spacek. 1992. *Measuring poles: A measuring device for surveying driving behavior on highways*. IVT-ETH Zurich and Planitronic Zurich, supported by the Swiss Road Safety Fund, Zurich, Switzerland.
11. Spacek, P. 1998. *Driving behavior and accident occurrence in curves driving behavior in curve areas*. Research Project 16/84, IVT-ETH, Swiss Federal Highways Office, Zurich, Switzerland.
12. Spacek, P. 2003. *Environmental impact of traffic, section: Road safety of traffic facilities*. Zurich, Switzerland: Lecture Notes, IVT-ETH.
13. Ren. Y.Y., Li, X.S., Ren, Y., Rachel, G., and W.W. Guo. 2011. Study on driving dangerous area in road curved section based on vehicle track characteristics. *International Journal of Computational Intelligence Systems* 4(6): 1237–1245.
14. Ren, Y.Y., Zhao, H.W., Li, X.S., and X.L. Zheng. 2012. Study on vehicle track model in road curved section based on vehicle dynamic characteristics. *Mathematical Problems in Engineering* ID 818136.
15. Donges, E. 1978. A two level model of driver steering behavior. *Human Factors* 20 (6): 691–707.
16. Beall, C., and J. Loomis. 1996. M. visual control of steering without course information. *Perception* 25: 481–494.

17. Kageyama, I., and H.B. Pacejka. 1992. On a new driver model with fuzzy control. *Vehicle System Dynamics: International Journal of Vehicle Mechanics and Mobility* 20 (Suppl): 314–324.
18. Muralikrishnan, K., K. Balasubramanian, S.M. Jawahar, et al. 2013. Poincare plot of heart rate variability: An approach towards explaning the cardiovascular autonomic function in obesity. *Indian Journal of Physiology and Pharmacology* 57 (1): 31–37.
19. Kwmen, P.W., H. Krum, and A.M. Tonkin. 1996. Poincare plot of heart rate variability allows quantitative display of parasympathetic nervous activity in humans. *Clinical Science* 91: 201–208.

Pedestrian Detection and Counting in Crowded Scenes

Juan Li, Qinglian He, Liya Yang and Chunfu Shao

Abstract Pedestrians are the most essential and important component of traffic systems. The pedestrian injury and fatality rates are at a high level due to the severe traffic crashes. Therefore, effective strategies should be implemented to enhance pedestrian safety. However, there is a lack of feasible methods to collect pedestrian data for pedestrian safety study. And the effectiveness of the existing methods may decrease along with the increasing complexity of the traffic system. To ensure pedestrian safety even in crowded scenes, a head-based pedestrian detection and counting method is proposed in this paper to capture the data of pedestrians. From the test results, several important attributes such as crowd density, location, and speed can be obtained. Instead of collecting the full bodies of pedestrians, human heads are used in our study to avoid the occlusion problem happened in crowded scenes. After setting the detection region, head detection is started by applying mixed color algorithm to locate candidate head area and then using Canny algorithm and Hough transform to extract target contour and locate head precisely. Finally, the minimum distance method is utilized to match and count the effective heads. The detection results compared with manual count indicate its extremely accurate performance. This method demonstrates the proposed approach which is useful and effective for crowded pedestrian detection and counting, and can be applied in real-world traffic system to detect pedestrians and prevent pedestrian accidents.

Keywords Pedestrian detection · Crowded scenes · Mixed color algorithm · Canny algorithm · Hough transform

J. Li (✉) · Q. He · C. Shao
MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology,
Beijing Jiaotong University, Beijing, China
e-mail: juanli@bjtu.edu.cn

L. Yang
School of Public Administration and Policy, Renmin University of China, Beijing, China

Introduction

Pedestrian safety is an important filed of traffic safety studies. With the development of the modernization and urbanization, the population becomes more and more concentrated. Large gatherings of people may cause severe crowd disasters in some special events. And in today's car-centered traffic system, the pedestrians' accident rate is at a high level because the pedestrians play a weak role in the traffic system. Several factors may influence pedestrian behavior and safety, such as environmental factors (roadway types, intersections, and traffic controls), traffic factors (volume, speed, and stationary vehicles), personal factors (physical abilities, experiences, and psychological states), and social factors (presence of others and trip purpose). Thus to ensure pedestrian safety, it is not only to study carefully from the characteristics of pedestrians, but also to coordinate the relationship between pedestrians and other components of the traffic system. It is urgent to change the car-centered traffic system into a people-centered system and ensure to make appropriate decisions for pedestrian safety improvements depending on the technology for the pedestrian detection and behavior analysis.

As the most important components of the traffic system, pedestrians should attract the most attention in the traffic studies. However, due to the great convenience brought from cars and the flexibility of pedestrians, most of the intelligent transportation applications are focused on vehicles, and few are extended to pedestrian studies. To better understand pedestrians' behaviors and provide them with a harmonious and safe traffic environment, effective approaches should be applied to accurately detect and locate pedestrians in scenes, especially in crowded scenes. After figuring out each individual's accurate location in scenes, the crowd density and location are obtained. Then, several meaningful applications can be put into practice, such as crowd management, visual surveillance, and public space design. The crowd management aims at avoiding crowd disasters when holding big events (large concerts and sport matches) and insuring public safety. And the visual surveillance is applied to detect anomalies and alarms and can help the public be fully prepared in advance. Besides, the pedestrian detection can also be used for public space design based on the crowd density. All the applications mentioned above are utilized to improve pedestrian safety and make appropriate decisions for emergency control, which indicate that pedestrian detection technology is of great importance in the Intelligent Traffic System and can enhance pedestrian safety. However, pedestrians' activities are usually free and have weak textures. The movements are hard to be detected by common sensors such as loops, sonar, and microwaves. Video cameras are widely used now for its competitive cost and easy operation. The intelligent video surveillance can be employed to detect pedestrians.

Endow machines with the ability to interact with people and recognize each person in the scene is one of the most important and potential challenges for the modern pedestrian accident prevention. There are two main approaches based on the intelligent video surveillance to detect pedestrian: direct and indirect methods [1]. The first one aims at identifying each individual in crowd scenes. This method

can be further classified into three approaches: part-based, model-based and trajectory-clustering-based approaches [2–4]. Direct pedestrian detection methods prefer to identify each pedestrian in the image sequences. Among these methods, part-based approaches contain the head detector, omega shape detector (head-shoulder detector), and the complete human shape detector. Lin et al. [5] used Haar Wavelet transform (HWT) to extract features of the human head contour and then utilized the support vector machines (SVM) to classify the head. However, the computational loading of this method is too heavy and may not function well when facing unclear head contours. Gall and Lempitsky [6] presented an improved Hough forests for tracking and recognition, which is robust to partial occlusions. Sidla et al. [7] created algorithm to count pedestrians in crowded scenes by detecting human head-shoulder regions (Ω-like shape), which is robust for individual tracking. Model-based approaches aim at tracking and extracting each individual in the image and then detecting and counting the number of pedestrians by using a shape-based model. These methods may have some difficulties in gathering features of all potential individuals. As for trajectory-clustering-based methods, they cluster all interest points on people who are being detected to track and determine their every motion in the crowded scene. Brostow et al. [8] applied an unsupervised Bayesian clustering framework to detect pedestrians in crowded scenes. However, when more than one people share the same trajectory, severe overlapping problems may occur. Compared to direct methods, the indirect crowd detection methods prefer extracting local and holistic features from the detected targets in foreground images [9]. Detecting features instead of the motion of people can be much easier and more accurate. Therefore, various features have been taken into account, such as pixel features, texture features, and corner-point features [8]. After collecting all this features, methods like linear regression function, Gaussian process, or neural networks are applied to detect and count the people in scenes [10–12]. Pixel-based approaches are more applicable to estimate the density of the crowd rather than identify each individual in the detected scene, only focus on the very local features. For example, Hussain et al. [13] proposed an automatic pixel-based model to detect the crowd density by using BPNN with extracted foreground blob pixels. But the detection accuracy of this method may decrease when facing high occlusion problems. Unlike pixel-based method, texture-based approaches mainly attempt to estimate the number of people in the scene. And corner-based approach proposed by Albiol et al. [14] uses the moving corner points, based on Harris algorithm [15] as features to detect and count the moving people in the scene. Without occlusion problems, these methods above prove that the relationship between the foreground area and the number of people in the detected scene is almost linear. Thus, it can be concluded that detecting pedestrians becomes much more complex when facing occlusion situations. And previous studies show that part-based detectors and the local feature collections can mitigate the occlusion problem [16].

In the urban roadway systems, for pedestrian safety analysis, each individual must be recognized in the scene. However, the crowded scenes increase the complexity of figuring out each person because there are multiple moving objects and

may have severe overlapped problems in the scene. Especially, it is hard to detect the complete shape of each individual pedestrian in crowded scenes; however, we can get the complete shape of human head easily. Head occlusion seldom occurs due to the length of shoulder. Therefore, this paper proposes a simple and effective head-based pedestrian detection method, which has high detection accuracy even in crowded scenes within high pedestrian density. And real-world data are adopted to verify the effectiveness of our new approach. The results show that the proposed method is feasible and effective. This method can detect pedestrian flows in all directions precisely without being affected by the pedestrians' density.

The rest of this paper is organized as follows: Section "Methodology" presents the proposed main methods used for pedestrian detection in crowded scenes. Section "Experimental Results" shows the experiments' results based on the real-world data. At last, Section "Conclusion" summarizes the study and discusses future works.

Methodology

The head-based detection method can be divided into three categories: template matching, statistical learning theory, and invariant features recognition [17–21]. And most methods are based on the characteristics of human head, such as the characteristics of face, hair, and skin. Therefore, the head-based pedestrian detection method proposed in this paper focuses on the color of pedestrians' hair and skin and features of human head contour.

The head-based pedestrian detection method proposed in this paper mainly contains three steps: detection region configuration, human head detection, and matching. Figure 1 presents a flowchart of our method. And the most important step is the human head detection. After setting the detection region precisely, the head detection procedure is started in the identified detection region by the mixed color algorithm to determine the candidate head region. Then, Canny algorithm is used to extract the target contour. For the next step, Hough transform is applied to locate head based on the shape of human head. At last, the minimum distance method is utilized to match and count the effective head in the detection region.
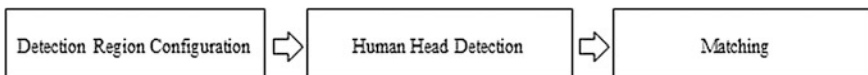
Fig. 1 Flowchart of the method

Detection Region Configuration

The detection region includes detection rectangles and lines. In this region, human heads can be detected within a certain range, and pedestrian flow can be counted in different directions precisely. Specifically, the following head detection process is implemented in detection rectangles, and the number of effective head is determined by the detected heads whether crossing over the detection lines. To improve the detection accuracy, the width of detection rectangles and lines should meet the following constraints as shown in Fig. 2. The gray lines present crosswalk, the purple line is the detection line, and the blue rectangle demonstrates the detection area.

I. Pedestrians who walk at the fastest speed should be detected twice by the detection line:

$$W_L > 2L - W \tag{1}$$

II. Pedestrians who walk at the lowest speed should not be detected by the detection line for three times:

$$W_L < 2L \tag{2}$$

III. The width of detection rectangles W_R should meet the following constraint:

$$W_R > W_L + 2W \tag{3}$$

L and W denote pedestrians' walking displacement and the width of peoples' heads. According to the survey of Beijing University of Technology [22], pedestrians' walking speed ranges from 1 to 1.5 m/s, and the average velocity is 1.2 m/s.
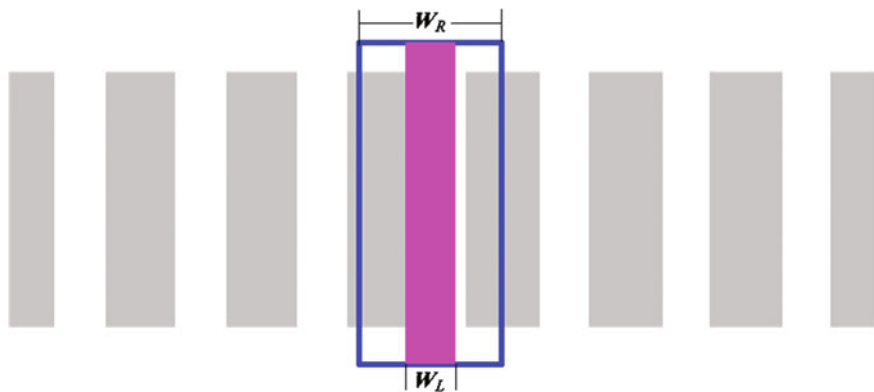
Fig. 2 Skeleton map of detection rectangle and line configuration

Therefore, this paper collects the digital image on the speed of 25 frames per second, and for every 5 frames we extract an image to handle. In another word, every time when an image is extracted, pedestrians' walking displacement ranges from 20 to 30 cm and the average displacement is 25 cm.

USA's Highway Capacity Manual (HCM 2000) [23] points out that the minimum space requirement of the pedestrian is determined by human height and shoulder width. As shown in Fig. 3, generally the pedestrian is considered as a 46 cm × 61 cm (thorax thickness B × shoulder width S) ellipse, which means that the total area of an individual is 2806 cm^2. And the top view of human head is usually regarded as a 30 cm × 30 cm (head width W × head width W) circle.

Based on the above conditions, the width of detection lines W_L can take the value of 30–60 cm, and the width of the corresponding detection rectangles can range between 100 and 130 cm. Besides, the width of detection line should be slightly longer than the width of the zebra crossing.

Human Head Detection

The color of pedestrians' hair and skin and features of human head contour are extracted from the image to identify the candidate head region. And the human head detection step is divided into three phases: candidate head region, head contour extraction, and head localization. Figure 4 shows the flowchart of head detection.

Candidate Head Region

In computer vision, images cannot be analyzed directly. They should be transformed into data for further studies, and the color space is the efficient way to represent the color [24]. Besides, the color space can be divided into two types

Fig. 4 Flowchart of head detection

Use the mixed color model to get the candidate head region

↓

Use the edge-detection algorithm to extract target contour

↓

Use Hough Transform to carry on head localization

↓

Carry on the head judgment based on the setting of hold value
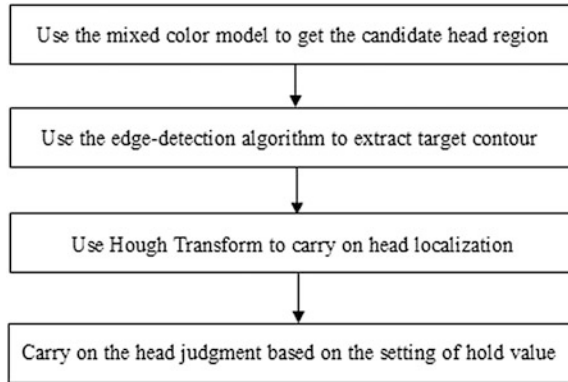
based on whether the representation results of the color using the selected color space may vary with the different display devices. Additionally, the results of the color space with different color representations when using different display devices is called a device-dependent color model, such as HSV and RGB. And the examples of the color space which are unrelated to the display devices are XYZ.

The color of Asians' hair is deep, mostly is black, and the difference of skin color shows in the brightness rather than the chroma. Therefore, the template of head images can be achieved by utilizing the clustering characteristics of the head and skin's color information. And to get the candidate head information more completely, mixed color algorithm is proposed in this paper to determine the candidate head region. First, establish color models in both RGB and HSV color space, respectively, and then get the mixed color model. At last, utilize the clustering features of the color of hair and skin in the mixed color model to get the candidate head area.

The RGB color space is an additive color space based on the RGB color model. It has three primary color components (red, green, and blue) and can produce any chromaticity that is the triangle defined by those primary colors. Use the RGB triplet (r, g, b) to represent the three-dimensional coordinate of the point of the given color by using the nonnegative values within a 0–1 range. Assign black to the origin at the vertex $(0, 0, 0)$ and with increasing intensity values running along the three axes up to white at the vertex $(1, 1, 1)$. As for the HSV color space, it also has three components (hue, saturation, and value) and it is more intuitive and perceptually relevant than RGB. Each color can be represented by the special HSV triplet (h, s, v) in the HSV cone. Specifically, the angle around the central vertical axis corresponds to "hue" with the value varies from 0° to 360°. The distance from the axis corresponds to "saturation" with the value ranges from 0 to 1. And the distance along the axis corresponds to "value" also with the value ranges from 0 to 1. The mixed color model is employed to get the candidate head region by gathering the clustering features in both RGB and HSV models, and the human head region can be obtained precisely. The flowchart of the candidate head region's determination is shown in Fig. 5. First, use the mixed color model of the hair to segment the

Fig. 5 Flowchart of determination of candidate head region
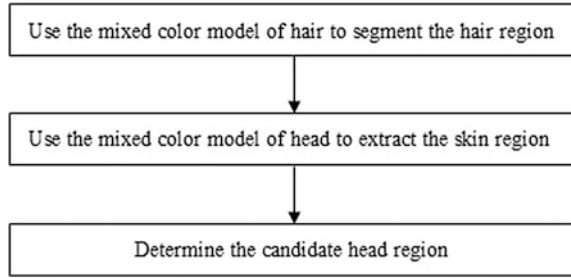
candidate hair region. To make the contour of the hair region much smoother, the mathematical morphology method (including closing operation and opening operation) is utilized to remove small pieces of interference and noise. Together with closing operation, opening operation removes small objects from the foreground, placing them in the background, while closing operation removes small holes in the foreground, changing small islands of background into foreground. Both of them remove the noise and also find specific shapes in the image. Then, the mixed color model of the human head is used to extract the candidate skin area and can easily remove the non-face region based on the principle that there must be some hair on the upper part of the face. Finally, the candidate head region is obtained by gathering the hair and skin regions.

The statistics analysis of the samples shows that the pixels in hair area with RGB color space are between 0 and 35 of the R component. After normalization, the interval [0–0.15] is acquired. By segmenting the image, the candidate hair area is obtained. Meanwhile, people's different skin color mainly reflected in brightness, that is, the information of (h, s) is independent with the value information (v) in the HSV color space. However, obvious clustering features are shown in the $H–S$ plane. The results show that, after the normalization of color pixels of both RGB and HSV models, the mixed color model of skin is obtained as follows:

$$0.005 < h < 0.10, 0.20 < s < 0.65, 0.60 < v < 1 \tag{4}$$

$$0.35 < r < 0.50, 0.25 < g < 0.35 \tag{5}$$

where h, s, and v mean hue, saturation, and value, respectively. And r and g represent the value of red and green light.

And to improve the detection precision, three aspects must be taken into consideration to remove the interference information of the candidate head region:

I. Remove the area of less than 50 pixels, which may be noise.
II. The head centroid must be in the connected region. The contour of the human head should be convex. Therefore, the non-head area with a concave shape can be removed by determining the position of the head centroid.
III. The length-to-width ratio of human head is within a certain limit: $0.7 \leq \text{width}/\text{height} \leq 1.5$

Finally, the candidate head regions are determined by gathering the above hair and skin regions, and the detection precision is improved by the interference removal methods.

Head Contour Extraction

The points, image brightness change sharply, are typically organized into a set of curved line segments termed edges. And the ideal results of detecting edges may lead to a set of connected curves that indicate the boundaries of the candidate objects. These results provide valuable events and changes in properties of the world to researchers. Therefore, applying an edge detection algorithm to an image not only significantly simplifies the analysis of image by reducing the amount of data to be processed, but also gives the available structural information about the boundaries of the target objects. The main purpose of this phase is to use the edge detection algorithm to extract target head contours in the candidate head region. Canny algorithm is adopted for its high signal-to-noise ratio (SNR) and detection precision.

As an optimization algorithm, Canny algorithm is based on three performance criteria raised by Canny [25]. The three principles are good detection, good localization, and only one response to a single edge. Due to its lower detection error rate and higher localization accuracy, Canny algorithm is recognized as the one of the most reliable and popular edge detection methods. To satisfy the above three strict requirements, Canny algorithm uses the first derivative of a Gaussian to compute the gradient and determines candidate edge points by finding the local maximum value of the gradient. Then, double-threshold method is utilized to detect sharp edges and weak edges. Only by connecting both of them can finally accurate the edge points. The detailed steps used in this study are as follows:

Step 1: Apply Gaussian filter to smooth the image in order to remove the noise. It is essential to filter out the noise to prevent false detection, and the Gaussian filter is applied to convolve with the image. The first derivative of a Gaussian function not only resists the noise but also provides the localization precisely; therefore, the first derivative of a two-dimensional Gaussian function is adopted in Canny algorithm to smooth the image.

Step 2: Use the finite difference of the first derivative to calculate the gradient's amplitude and direction. And the template of the first derivative convolution is:

$$H_x = \begin{vmatrix} -1 & -1 \\ 1 & 1 \end{vmatrix}, \quad H_y = \begin{vmatrix} 1 & -1 \\ 1 & -1 \end{vmatrix} \tag{6}$$

The derivatives of three channels of x direction: R_x, G_x, and B_x and their combination P_x are shown as follows:

$$
\begin{aligned}
R_x &= [f_R(i,j+1) + f_R(i+1,j+1)] - [f_R(i,j) + f_R(i+1,j)] \\
G_x &= [f_G(i,j+1) + f_G(i+1,j+1)] - [f_G(i,j) + f_G(i+1,j)] \\
B_x &= [f_B(i,j+1) + f_B(i+1,j+1)] - [f_B(i,j) + f_B(i+1,j)] \\
P_x &= R_x + G_x + B_x
\end{aligned}
\tag{7}
$$

where $f_R(i, j)$, $f_G(i, j)$, and $f_B(i, j)$ represent the pixel values of the three channels of the image. Then, the amplitude (M) and direction (q) of the gradient can be expressed as:

$$
M = M_R + M_G + M_B = \sqrt{R_x^2 + R_y^2} + \sqrt{G_x^2 + G_y^2} + \sqrt{B_x^2 + B_y^2}
\tag{8}
$$

$$
q = arctan\,(p_y/p_x)
\tag{9}
$$

Step 3: Create the refined edges by using non-maximum suppression (NMS) method. After the two steps above, the edge determined by the gradient calculation is still blurred. According to criterion 3, there should be only one response to a single edge. Therefore, non-maximum suppression is applied to guarantee that criterion by suppressing all the gradient values to zero except the local maximal. Edges should be determined by both holistic and local pixels. The NMS method compares the amplitude of the center pixel in the neighborhood with the two adjacent pixels along the gradient direction. If the amplitude of center pixel is the largest compared to the other two pixels, the value will be preserved. Otherwise, the value will be suppressed to zero.

Step 4: Detect and connect the defined edges by using double-threshold method. The edge pixels determined by NMS method are quite accurate to present the real edge. However, it still contains some fake edge pixels caused by noise and small textures, thus the double-threshold method is proposed to solve this problem. Two threshold values are set to clarify the different types of edge pixels: τ_1 and τ_2, and let $\tau_1 = 2\tau_2$ to get two corresponding edge images T_1 and T_2. The edge image obtained from the high threshold T_2 has some broken edge contours due to the numerous fake errors. The results of the image from the low threshold T_1 are just the opposite. Therefore, the image from the low threshold T_1 can be the supplement of the image from the high threshold T_2, that is, it can connect broken edge contours in T_2 and then get the final candidate human head contour.

Head Localization

In the image sequence, the shape of human head is like the ellipse. Therefore, Hough transform is utilized to locate the human head in the target contour.

Hough transform method is mainly used to detect lines, curves, and ellipses by performing an explicit voting procedure of several parameterized image objects to

assign the edge points to the object candidates [26]. If the edge points meet the requirement, the shape of the candidates is then determined.

Hough transform cannot be applied to locate human head directly, for the edge of human head is an irregular ellipse with curvy and unclosed contour. Besides, most edge pixels of head are not on the desired ellipse contour. Thus, the strategy to locate human head precisely is to find an ellipse roughly coincided with the head and make it contained within a ring, and then Hough transform can be applied to detect that ring in order to achieve the accurate head localization. As shown in Fig. 6, the inner and outer circles are concentric, and the radii of the two circles are the ellipse's minor semi-axis and major semi-axis, respectively. The location and size of the ring are determined by the circle center (a_0, b_0) and the radius of the inner circle r_0. Assume that the radius ratio of the inner circle and the outer one is $1:n$, where the value of n depends on the camera configuration and the value of n is 1.2 in overlooking state. Therefore, the contour of human head is contained in a special ring, whose radii of the inner and outer circle are r_0 and $1.2r_0$, respectively. Based on the size of the detected head and the detection precision, the minimum radius of the inner circle r is set as 5 in our study. Then, the radius of the inner circle r_0 varies from r to $1.2r$. And the detailed steps are as follows:

Step 1: Estimate the center position of the head based on the boundary rectangle of the head contour.
Step 2: Hough transform is utilized to find the accurate center position (a_0, b_0) in the boundary rectangle.
Step 3: Take the (a_0, b_0) as the circle center to carry on Hough transform with the radius of inner circle varies from r to $1.2r$.
Step 4: Choose the maximum value of all the experiments' results to be the final selected radius.

At last, the accurate head contour can be determined by the circle center (a_0, b_0) and the certain ring (the radius of the inner and outer circle is r_0 and $1.2r_0$).

Fig. 6 Skeleton map of head shape
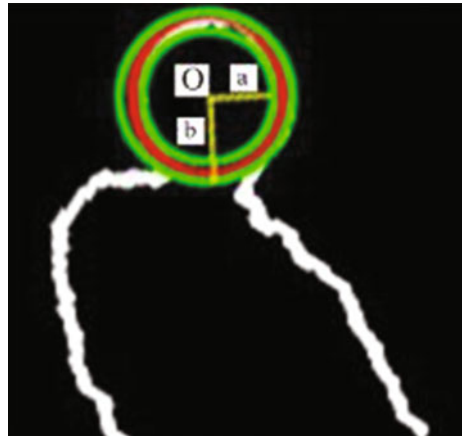
To make the results of head contour detection more accurate, two constraints are adopted to examine the detected ring: (1) the inner circle area $A_{incircl}$ should vary from A_{min} to A_{max}. And the value of A_{min} and A_{max} is related to the camera's location and observation range. (2) The ratio of the boundary rectangle of head A_{head} to the outer circle $A_{excircled}$ should meet the constraint: $1 < A_{head}/A_{excircled} < 1.3$. If both conditions are met, the contours of heads are proved to be accurate.

Matching

A matching method is proposed in our study by using direction judgment record. This record is a link list for saving the effective head information collected at the first detection. If the following detection's result matches with the former one, judge the direction of its displacement. Positive displacement represents the forward direction; otherwise, it should be the negative direction. And if there is no match, just delete it as interference information. Detailed steps are as follows and Fig. 7 shows the flowchart of head matching process.
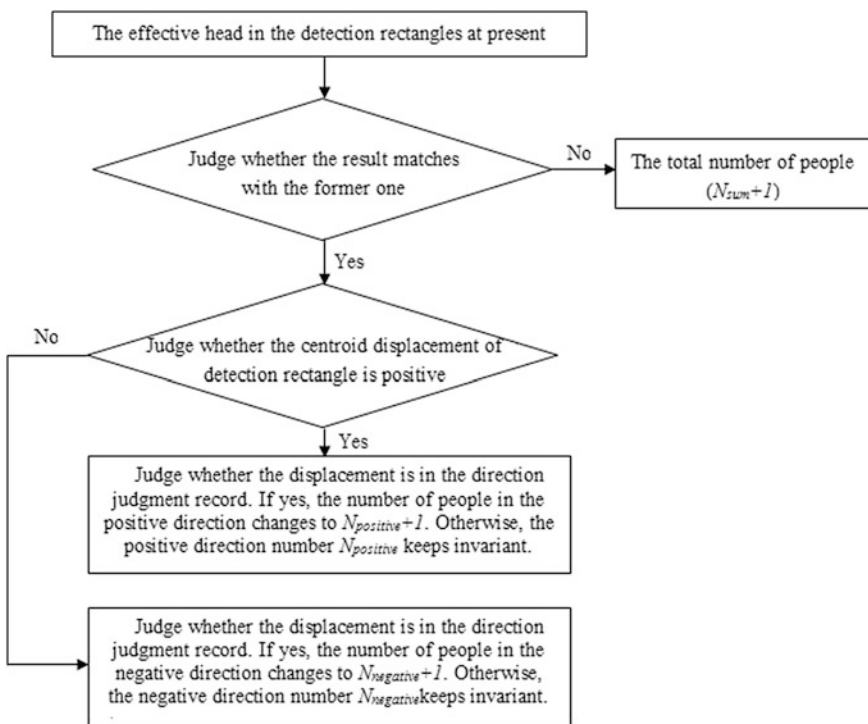
Fig. 7 Flowchart of head match and counting

Step 1: Input the effective head in the detection rectangles at present.

Step 2: Judge whether the result matches with the former one. If no, go straight to the next step. Otherwise, turn to Step 4.

Step 3: Put the total number of people (N_{sum} + 1) into direction judgment record.

Step 4: Judge whether the centroid displacement of detection rectangle is positive. If yes, go straight to the next step. Otherwise, turn to Step 6.

Step 5: Judge whether the displacement is in the direction judgment record. If yes, the number of people in the positive direction changes to $N_{positive}$ + 1. Otherwise, the positive direction number $N_{positive}$ keeps invariant.

Step 6: Judge whether the displacement is in the direction judgment record. If yes, the number of people in the negative direction changes to $N_{negative}$ + 1. Otherwise, the negative direction number $N_{negative}$ keeps invariant.

It should be noted that the positive direction is from east to west and the negative direction is the opposite in this study. Besides, the effective head matching process is achieved by judging the head distance between the current image and the former one. If that distance is less than thickness of thorax B and shoulder width S and maximum walking distance L_{max}, we can use the minimum distance method for matching.

Experimental Results

To verify the effectiveness of the head-based pedestrian detection method proposed in this paper, the practical traffic video is applied. The video sequence using in this paper was collected on March 18, 2010 at the crosswalk in front of the east gate of Beijing Jiaotong University. The format of the file is AVI, and the video resolution is 720×576. Besides, the rate of the video is 25 frames per second, and the test site is shown in Fig. 8.

Fig. 8 Snapshot of the test site
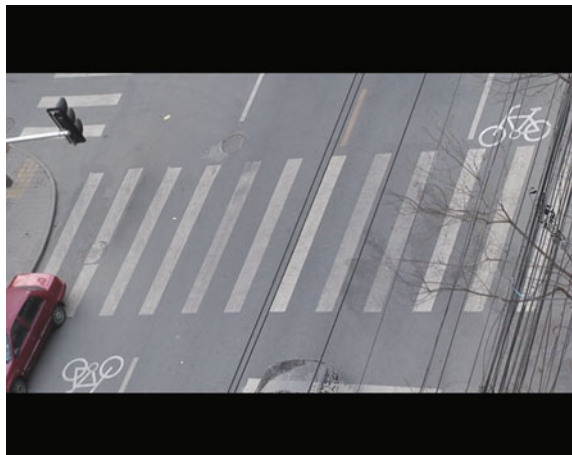
Detection rectangles (width 130 cm × length 550 cm) and lines (width 45 cm × length 550 cm) are set in the crosswalk. By detecting head in detection rectangles, we counted pedestrians in two directions (positive and negative). Specifically, the positive direction is from east to west and the negative is from west to east. The results of a certain moment in the detection process are shown in Fig. 9. Blue rectangle and purple line represent the detection rectangle and line, respectively, and the green elliptic cycle represents the detecting head in the detection rectangles.

It shown from Fig. 9 that the total number of detected pedestrians is 262, and 149 of them are in the positive direction while others are in the negative direction in the experiment. The average processing speed of the pedestrian detection is approximately 38 ms/frame, which indicates that it can meet the real-time requirements with good effect. And the results shown in Table 1 indicate that the head-based pedestrian detection and counting method proposed in this study not only avoids the occlusion problem, but also improves the detection precision. The number of pedestrians in the positive direction is 218, and 189 of them were accurately detected. It can be concluded that the correct detection rate of the positive direction is 87%. And the pedestrian numbers in the negative direction is 144 with 136 of them were accurately detected. The correct detection rate of the negative direction is 95%, which is a bit higher than that of the positive direction.

Fig. 9 Snapshot of the test result

Table 1 Head-based pedestrian detection result

Statistics	Pedestrian numbers	Detection numbers	Correct rate (%)
Positive direction	218	189	87
Negative direction	144	136	95
Total	362	325	90

Besides the flow rate of pedestrian can also be figured out, specifically is 1.21 people per second. This rate indicates that the traffic congestion state is crowded. And the overall correct detection rate is 90% (325 of the total 362 pedestrians were successfully detected). Therefore, the overall correct detection rate (90%) with the crowded road condition indicates that our head-based detection method is effective and can avoid the occlusion problems even in crowded scenes.

Conclusion

In this paper, we propose a novel and effective head-based pedestrian detection method, which has high detection accuracy even in crowded scenes within high people density. It contains three essential steps: detection region configuration, human head detection, and matching. As for implementing the most important step, the human head detection, mixed color algorithm is applied to determine the candidate head region. After that, Canny algorithm is utilized to extract the target contour. And Hough transform is used to achieve accurate head localization by detecting the certain ring. At last, matching step by using direction judgment record combined with the minimum distance method can detect and count the heads in the detection region.

The test results based on the real-world data show that the head-based pedestrian detection method proposed in this paper has a high correct detection rate (over 90%), which indicates that our method improves the detection precision and can be applied to solve the real-world pedestrian detection work effectively. Besides, instead of collecting complete shapes of pedestrians, our head-based pedestrian detection method can be used to efficiently detect pedestrians even in crowded scenes within high people density. Our method absolutely can solve the occlusion problem successfully and improves the state-of-the-art pedestrian detection methods. Above all, our method can be applied in real-world traffic system to prevent pedestrian accident and improve pedestrian safety.

As future work, we would like to expend the situations of test sites and deeply analyze the characteristics of pedestrian behaviors for a multitude of occluding accessories such as backpacks, and hats may perturb a pedestrian's head silhouette. Besides, further study would focus on depth understanding pedestrian behavior by recognizing their activities and forecasting potential traffic and pedestrian collisions.

References

1. Conte, D., P. Foggia, and G. Percannella. 2010. A method for counting people in crowded scenes. In *Proceedings of IEEE Advanced Video and Signal Based Surveillance (AVSS)*, 225–232.
2. Khatoon, R., S.M. Saqlain, and S. Bibi. 2012. A robust and enhanced approach for human detection in crowd. In *Proceedings of IEEE Multitopic Conference (INMIC)*, 215–221.
3. Zhan, B., D.N. Monekosso, and P. Remagnino. 2008. Crowd analysis: A survey. *Machine Vision and Applications* 19: 345–357.
4. Brostow, G.J., and R. Cipolla. 2006. Unsupervised bayesian detection of independent motion in crowds. In *Proceedings of IEEE Computer Vision and Pattern Recognition*, 594–601.
5. Lin, S.F., J.Y. Chen, and H.X. Chao. 2001. Estimation of number of people in crowded scenes using perspective transformation. *IEEE Transactions on Systems, Man, and Cybernetics Part A: Systems and Humans* 31 (6): 645–654.
6. Gall, J., and V. Lempitsky. 2013. Class-specific Hough forests for object detection. In *Decision Forests for Computer Vision and Medical Image Analysis*, 143–157. Springer.
7. Sidla, O., Y. Lypetskyy, N. Brandle, and S. Seer. 2006. Pedestrian detection and tracking for counting applications in crowded situations. In *Proceedings of IEEE Advanced Video and Signal Based Surveillance (AVSS)*.
8. Brostow, G.J., and R. Cipolla. 2006. Unsupervised Bayesian detection of independent motion in crowds. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 37–44.
9. Saleh, S.A.M., and H. Ibrahim. 2015. Recent survey on crowd density estimation and counting for visual surveillance. *Engineering Applications of Artificial Intelligence* 41: 103–114.
10. Davies, A.C., J.H. Yin, and S.A. Velastin. 1995. Crowd monitoring using image processing. *Electronics & Communication Engineering Journal* 7 (1): 37–47.
11. Chan, A.B., Z.S.J. Liang, and N. Vasconcelos. 2008. Privacy preserving crowd monitoring: Counting people without people models or tracking. In *CVPR, IEEE Computer Society Conference on Computer Vision and Pattern Recognition*, 1–7.
12. Hou, Y.L., and G.K.H. Pang. 2011. People counting and human detection in a challenging situation. *IEEE Transactions on Systems Man and Cybernetics—Part A Systems and Humans* 41(1): 24–33.
13. Hussain, N., H.S.M. Yatim, N.L. Hussain, et al. 2011. A pixel-based crowd density estimation system for Masjid al-Haram. *Safety Science* 49 (6): 824–833.
14. Albiol, A., J. María, A. Silla, J.M. Albiol. 2009. Video analysis using corner motion statistics. In *Proceedings of the IEEE International workshop on Performance Evaluation of Tracking & Surveillance—38 Tools Appl*, 31–38.
15. Harris, B.C., and M.A. Stephens. 2010. Combined corner and edge detector. In *Proceedings of the 4 Alvey Vision Conference*.
16. Leibe, B., E. Seemann, and B. Schiele. 2005. Pedestrian detection in crowded scenes. In *Proceedings of IEEE Computer Vision and Pattern Recognition*, 878–885.
17. Yang, M.H., D.J. Kriegman, and N. Ahuja. 2002. Detecting faces in images: A survey. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 24 (1): 34–58.
18. Mckenna, S.J., Y. Raja, and S. Gong. 1999. Tracking colour objects using adaptive mixture models. *Image and Vision Computing* 17 (3–4): 225–231.
19. Dai, Y., and Y. Nakano. 1996. Face-texture model based on SGLD and its application in face detection in a color scene. *Pattern Recognition* 29 (6): 1007–1017.
20. Miao, J., B. Yin, K. Wang, et al. 1999. A hierarchical multiscale and multiangle system for human face detection in a complex background using gravity-center template. *Pattern Recognition* 32 (7): 1237–1248.

21. Edwards, G.J., C.J. Taylor, and T.F. Cootes. 1998. Learning to identify and track faces in image sequences. *IEEE Computer Society* 260.
22. Sun, Z.Y. 2004. *Traffic characteristics of pedestrians at crosswalks of signalized intersections*. Beijing: Beijing University of Technology.
23. TRB. 2000. *Highway capacity manual (HCM 2000)*. Washington, DC: Transportation Research Board, National Research Council.
24. Ruan, Q. 2007. *Digital Image Processing Study*. Beijing: Electronic Industry Press.
25. Canny, J. 1986. A computational approach to edge detection. In *Proceedings of IEEE Pattern Analysis and Machine Intelligence*, 679–698.
26. Hough, P.V.C. 1962. Method and means for recognizing complex patterns. U.S. Patent 3069654.

Coordinated Carpool Route Selection and Optimization for Dynamic Uncertain Demand Based on Connected Vehicles

Guiliang Zhou, Tianwen Bao, Lina Mao, Lijun Huang and Dingxin Wu

Abstract The paper puts forward a method for selecting and optimizing coordinated carpool routes to meet the dynamic and uncertain demand based on connected vehicles. The distance, cost, and time of the selected route are calculated using Dijkstra's algorithm and distance matrix method combined with the cost function and road resistance function. Based on this, the concept of sample weight coefficient is introduced. Then, the model of sample weights for selecting the coordinated route is constructed. The results show that compared with a common route, the optimal coordinated carpool route could save the total distance by 38.46%, total cost by 17.66%, and total time by 21.57% under dynamic and uncertain demand. Application in Huai'an, Jiangsu shows that the coordinated carpool route can greatly reduce the cost and time.

G. Zhou (✉) · T. Bao · L. Mao · L. Huang · D. Wu
Faculty of Transportation Engineering, Huaiyin Institute of Technology,
Huaian 223003, China
e-mail: ZGLpaper@foxmail.com

L. Mao · D. Wu
School of Transportation, Southeast University, Nanjing 210096, China

G. Zhou
School of Automotive and Traffic Engineering, Jiangsu University,
Zhenjiang 212013, China

T. Bao
School of Transportation and Logistics, Southwest Jiaotong University,
Chengdu 610036, China

L. Huang
Faculty of Transportation, Shanghai Maritime University, Shanghai 201306, China

Introduction

In order to alleviate traffic pressure, improve the utilization rate of the private car, and reduce the environmental pollution, carpool has become a trend which cannot be halted. However, there are many uncertain factors with its development. The problems such as encountering random carpool and burst block waste riders' cost and time and reduce the users' experience. Focusing on these problems, the paper introduces solutions combining with the connected vehicle technology. Connected vehicles can collect information on roads, vehicles, and other traffic facilities randomly. Now the private cars are mainly shared by two riders. Once if more than two people participate in carpool, then there will be an extra increase in time and distance. It will lead to unreasonable carpool route for many riders. With the help of the traffic information acquisition through connected vehicles, we can deal with the carpool demand with dynamic uncertainty at any time by searching the optimal coordinated carpool route and realizing the coordinated carpool of many vehicles to save distance and time.

Review on Related Researches

The domestic and foreign scholars have studied the route optimization algorithm according to different demands. In the study of vehicle dynamic route guidance under the condition of vehicle networking, Xie [1] introduced a dynamic path algorithm in an uncertain environment combined with vehicle networking technology to get relevant traffic data. He put forward a dynamic route guidance method for dynamic route guidance based on vehicle steering delay. This approach is a benefit to prevent traffic jam and improve traffic network condition. It adapts to the practical application of transportation. Lin [2] introduced the genetic neural network into the dynamic shortest path algorithm. It not only optimized the network structure but also provided a new method to predict the traffic flow of the nonlinear network with uncertainty. She also added and improved the original static Dijkstra's algorithm to make the shortest path algorithm less complex. Basing on user's subjective demand in route optimization algorithm, Chen and Ruihua [3] discussed about the route choices with the passengers' subjective preference. When there is no subjective preference, the entropy theory of information and comprehensive decision method of multiple attributes could be used to get the algorithm model of the synthetic route attribute value. Otherwise, the subjective preference could be obtained through the reciprocal judgment matrix according to different routes, and then the route selection could be humanized designed. In coordinated transportation, most of the researches focus on minimizing the cost of the whole transportation process. Weng and Zhu [4] studied the coordination and integration of the transport routes. They set all the O-D pairs (i.e., the tasks of transport) in the specified length of the line to take any direct moving. Then, single point of transit

and the integration of two-point transit routes were planned. After that, intermediate nodes paid certain cost to produce a shipping discount, which helped selecting the integrated route of O-D pairs to minimize the total cost. Weng [5] worked on the coordination of complicated tasks for the optimal hub routes problem. It aimed at reaching the destination through 0, 1, or 2 hubs at a finite distance with cost discount paid to the fixed cost after the arc of hub line. Then, reduction in the inherent costs to the maximum extent would be realized. In terms of reducing the travel time of the entire network, a new path planning algorithm for network searching was proposed by Oh [6]. Suppose that the multiple autonomous vehicles are able to access each path efficiently, then an optimal solution to solve mixed integer linear programming could be found, which would achieve maximum reduction in travel time.

Many scholars involved in the route optimization algorithm mainly focus on the distance, time, and cost. A single algorithm on the three directions has been well developed [7]. This paper, based on existing researches, combines the route optimization with the real-time carpool and the route coordination. It calculates the distance, cost, and time of the selected route by Dijkstra's algorithm, gets the distance matrix combined with the cost function method and the road resistance function method. Then, the concept of sample weight coefficient is introduced to construct the sample weight calculation model, and the optimal selection of the coordinated route is carried out.

Algorithm on Coordinated Carpool Route Selection with Dynamic Uncertain Demand

With car owners' information released among connected vehicles, the route could be recommended by the system or determined by independent choice. When encountering two passengers on the way releasing information, it is needed to deviate from the original route. Then, the owner needs to select an optimal route or a few coordinated routes to get the passengers. According to the optimal principle of a dynamic system [8], the algorithm has the following four categories: ① Searching the shortest route, calculate the shortest distance from the start point to the end point; ② Setting time consumption as the research objective, calculate the average time consumption from beginning to the end; ③ Seeking the minimum cost, calculate the average cost or relative minimum cost for the whole carpool process; ④ Comprehensively considering the distance, time, and cost factors, use the weighting method. In this paper, we use the Dijkstra's algorithm, the distance matrix method, and the resistance function method to study the path distance, time, and cost. It combines the distance, time, and cost of the path by deciding the weights to get a comprehensive consideration of the optimal target [9].

First of all, according to the survey, weights set for the distance, time, and cost are α, β, and γ, respectively. Then start screening each path suitable for traveling.

Travel time prediction: using resistance function, we can collect many data from each section of the road at any time through connected vehicles. These data include the speed of a vehicle on the road at a certain time, the traveling speed of free flow, the density of traffic flow at a certain time, and traffic density during a traffic jam. Integrating these data with the survey data, the traffic capacity of the time section of the road could be calculated. Next, we can get the travel time of the road. The shortest path can be worked out using the Dijkstra's algorithm combined with the distance matrix algorithm. The cost function is built according to the shortest distance by calculating the cost of the whole process. Using Dijkstra's algorithm, we can get the shortest route and the most expensive route from the starting point to end point. The distance matrix algorithm is used to get the distance between any two nodes conveniently and accurately and make it easy to calculate the cost of each route. Meanwhile, we use the resistance function method to obtain the travel time. Then, weights are assigned to carry out a comprehensive calculation. It will not only suggest the best route to the customer but also recommend the routes with shortest time and distance, and the minimum cost. In order to fulfill different requirements of passengers, this recommendation is more humane.

Solve the Distance and Cost by Dijkstra's Algorithm and Distance Matrix Algorithm

Dijkstra's algorithm is also called label method. Firstly, a label is made for each vertex from the starting point. These labels are classified into two categories: T and P. The T label represents the upper bound of the shortest path from the starting point to this point; the P label represents the shortest path from the starting point to the point, which is called fixed label. The point which has been given the P label will be no longer changed, and the point where there is no labeled P is marked T. Each step of the algorithm will modify a point of T label into P label. After finite steps, all of the points would be marked P label and will get every point the right way from the starting point. The calculation steps are as follows:

1. Mark on the P label to the starting point and to make $P(1) = 0$. The rest of the points are marked on the T label. $T_0(j) = +\infty$ represents that the shortest path is 0 and the upper bound on the shortest path from the starting point to a point is infinity. The number in parentheses is the period, and the 0 corner is the initial value of the label.
2. Set A_i as the first round label (the first $N - 1$ wheel label) by just obtaining the points with P label and carry out a new round of labeling (section N) for points that do not have a P label. Suppose that all points A_j are adjacent to A_i and are not marked with the P label. Modify the T label of A_j as $T_N(j) = \min(T_j, P(i) + d_{ij})$

In the formula,

d_{ij} the distance from A_i to A_j

T_j T label already obtained at A_j points before N round mark

For all the T labels, the method to find a minimum T label $T_{N(j_0)}$ is as follows:

$$T_{N(j_0)} = \min(T_N(j), T(l))$$

In the formula,

$T(l)$ T label already obtained with A_1 adjacent A_i

Labeled T_{j_0} for the P label, that is $P(j_0) = T_N(j_0)$.

3. if there is no T label in the formula, the algorithm ends. Otherwise go to Step 2.

Solve the Distance Between Any Two Nodes and Construct the Cost Function by Distance Matrix

Distance Matrix

The adjacency matrix representation provides the following formula for representing the connection information of the two nodes.

$$A[i,j] = \begin{cases} 0 & \text{Coincident node}: i \text{ and } j \\ \infty & \text{Nodes are not connected}: i \text{ and } j \\ w & \text{Node } i \text{ and } j \text{ connection on the weight of } W \end{cases}$$

Among them, $A[i, j]$ stands for the distance between i and j in the adjacency matrix. Many matrices such as $D^{(1)}, D^{(2)}, \ldots, D^{(n)}$ are constructed by using the method of inserting vertices in the weighted adjacency matrix of graphs, respectively. The matrix of the graph is obtained at last. Besides, the insertion point matrix is also obtained to get the shortest path between two points.

Algorithm principle: The weighted adjacency matrix is used as the initial value of the distance matrix.

1. $D^{(1)} = (d_{ij}^{(1)})_{n \times n}(i, j = 1, 2, \ldots, n)$.
2. In $D^{(n)}$, if $i = j$, $d_{ij}^{n+1} = 0$, otherwise, turn to (3).
3. If $d_{il}^k = \infty$ or $d_{lj}^k = \infty(1 = 1, 2, 3, \ldots, n)$, and $d_{ij}^k \neq \infty$, $d_{ij}^{k+1} = d_{ij}^k$; otherwise, turn to (4).

4. If $\quad i < j, \quad d_{ij}^{n+1} = \min \quad \left\{ \min \left\{ \min_{l<i} \left\{ d_{il}^{k+1} + d_{lj}^{k+1} \right\}, \min_{i \leq l \leq j} \left\{ d_{il}^{k+1} + d_{lj}^{k} \right\}, \right. \right.$

$\left. \left. \min_{j \leq l} \left\{ d_{il}^{k} + d_{lj}^{k} \right\} \right\} \right\}; \qquad$ otherwise, $\qquad d_{ij}^{k+1} = \min \left\{ \min_{l<i} \left\{ d_{il}^{k+1} + d_{lj}^{k+1} \right\}, \right.$

$\left. \min_{i \leq l < j} \left\{ d_{ij}^{k+1} + d_{ij}^{k} \right\}, \min_{j \leq l} \left\{ d_{il}^{k} + d_{lj}^{k} \right\} \right.$

If $d_{ij}^{k+1} = d_{il}^{k} + d_{lj}^{k}$, and $d_{ij}^{k+1} < d_{ij}^{k}$, then write down 1 and the corresponding elements in the distance matrix are replaced by 1. It shows that the length of the shortest path from node v_i to node v_j after the iteration is d_{ij}^{k+1}.

5. If $D^{(k+1)} = D^{(k)}$, ends here; otherwise, $k = k+1$, turn to (2).

Calculate the Total Cost of Unit Distance of the Route

The optimal route system recommends the route before the carpool. But the new customers may increase dynamically in the actual carpool process, and the route and cost will change accordingly. Changes in the number of customers will lead to changes in the cost for each customer. More customers participate in carpool will lead to lower unit cost for customers. Optimal route and unit distance costs of the route will be changed after real-time carpool. Therefore, we need to reanalyze the entire network map. In order to optimize the algorithm, the paper only considers the influence of distance constraint. So the total cost of the unit distance of the route can be calculated. The formula is as follows:

$$f_{ij}^{km} = d_{ik} \times q_{ik} + d_{km} \times q_{km} + d_{mj} \times q_{mj}$$

f_{ij}^{km} is the total cost of unit distance of the route $i \rightarrow k \rightarrow m \rightarrow j$. d_{ij} represents the shortest distance of $i \rightarrow j$. d_{km} represents the shortest distance of $k \rightarrow m$. q_{ij} represents unit distance of the route of $i \rightarrow j$, and q_{km} represents the unit distance of the route of $k \rightarrow m$.

Calculate the Travel Time Using Resistance Function

Most customers' requirements on the travel time are relatively high in the carpool process. Once the shortest travel time is the optimal objective, then it is necessary to analyze the travel time of the routes [10]. Generally, the average time of the road is taken as the weight. A simple algorithm is used to calculate the ratio between the distance and the design speed. However, this scheme cannot reflect the characteristics of traffic flow and the reality. It is not suitable for real-time changes in the carpool process. Besides, congestion, accidents, and other things may happen. So

resistance function is introduced to determine the cost of road resistance and carpool time. Any two nodes in the network graph are regarded as an OD pair, and there are m random lines between each OD pair. Each line has its own resistance and capacity. Taking the potential obstruction of the selected path into account, which will lead to changes in the resistance of the road. Then, the congestion coefficient increases and the capacity decreases, which will take a lot of time.

Construct the formula of road resistance function as follows:

$$v_k = v_f - \frac{v_f}{k_{max}} \times k$$

In the formula,

v_k The speed of a vehicle on a road at a certain time
v_f Traveling speed of free flow
k The density of traffic flow on a road at a certain time
k_{max} Traffic density in a traffic jam

According to the relation $q_k = v_k \times k$, in which q_k is the traffic flow, we can deduce the formula

$$q_k = v_f \times k - \frac{v_f}{k_{max}} \times k^2$$

The maximum traffic flow can be derived from the formula when $k = \frac{k_{max}}{2}$.

That is $q_{max} = \frac{v_f \times k_{max}}{4}$, which is also the capacity of this section C. Take $k = \frac{k_{max}}{2}$ into the resistance function formula, the speed can be obtained as $v_k = \frac{v_f}{2}$. Assume that the length of the section is l, then the travel time of the road is $t = \frac{l}{v_k}$. For the free running state, the travel time is $t_0 = \frac{l}{v_f}$. According to t and t_0, we can obtain the speed $v_k = \frac{t_0}{t} \times v_f$. The resistance function can be expressed as a function of k, that is $\frac{k_{max}}{v_f}(v_f - v_k)$; taking v_k and the function of k into q_k, we can obtain:

$$q_k = -\frac{k_{max}}{v_f} v_k^2 + k_{max} v_k = -C\left(\frac{t_0}{t}\right)^2 + C\frac{t_0}{t}$$

It can be converted into the equation in the form of $\frac{t_0}{t}$, that is $\left(\frac{t_0}{t}\right)^2 - \frac{t_0}{t} + \frac{q_k}{C} = 0$.

Suppose no accident occurred, namely $q_k \leq C$, then we can obtain: $\frac{t_0}{t} = \frac{1}{2}\left(1 \pm \sqrt{1 - \frac{q}{C}}\right)$.

Thus, the travel time formula is: $\dfrac{2t_0}{1 \pm \sqrt{1 - \frac{q}{C}}} = \dfrac{2t_0}{1 \pm \sqrt{1 - \frac{q}{\frac{v_f k_{max}}{4}}}}$.

Comprehensive Weight Calculation

In this paper, we need to decide, as well as use, a certain algorithm to calculate the weights on the distance, time, and cost of the route. However, due to different riders have different focuses on the three factors, different weight distribution scheme will have great influence on the results. Then, we introduce the sample weight coefficient, which is defined as: $F_i = \alpha d + \beta t + \gamma c$.

In the formula, i is the route and $i > 1$. α, β, and γ represent the weight of the distance, time, and cost of the route, respectively. d, t, and c represent the distance, time, and cost of each route, respectively. In order to make comparison, it should be considered that the systems of units of distance, time, and cost are different, which should be normalized [11]. The min-max standard is selected here, and the data is normalized to [0, 1] range with the conversion function: $x_k = \frac{x - x_{\min}}{x_{\max} - x_{\min}}$.

In the formula, x is the sample data. x_k is the normalized sample data. $x_{\min}$ is the minimum value of the sample data. $x_{\max}$ is the maximum value of the sample data. The sample weight coefficients are combined with the variation of distance, time, and cost weight and their changes. The smaller the sample weight coefficient is, the better the selected route will be.

Selection and Optimization of Coordinated Carpool Route Based on Dynamic and Uncertain Demand—a Case Study of Huai'an, Jiangsu

Traffic congestion is becoming increasingly serious in Huai'an, Jiangsu province, and private carpool has become one of the commute modes. The key laboratory of transportation and security of Jiangsu Province has studied on private carpool savings in Huai'an based on connected vehicles. Considering the high complexity of the whole road network [12], this paper only selects the partial network of South Chengde Road in Huai'an. The model of the routes is shown in Fig. 1.

It is can be seen from Fig. 1 that the blue line is the fixed carpool route at the beginning. It is set by the system of connected vehicles. A is the starting point, B is the deviation point, D is a dynamic non-fixed point, K is the middle of the cooperative node, M is the end point, and L is the point of the final destination for dynamic uncertain demand. If the driver decided to carry the passenger who later joined during real-time carpool, the driver is likely to deviate from the original route. Then, the final destination is different between the original customers and the later passengers. If we sent different riders to different route destinations, it will consume more time and cost, and it is impossible to realize system optimization on multi-route design. Therefore, the selection of coordinated carpool route becomes extremely important. L is known as the final destination of real-time carpool customers in Fig. 1, and K is the cooperative node. KL is the optimal route between point K and point L. In order to optimize the coordinated carpool, we can search the
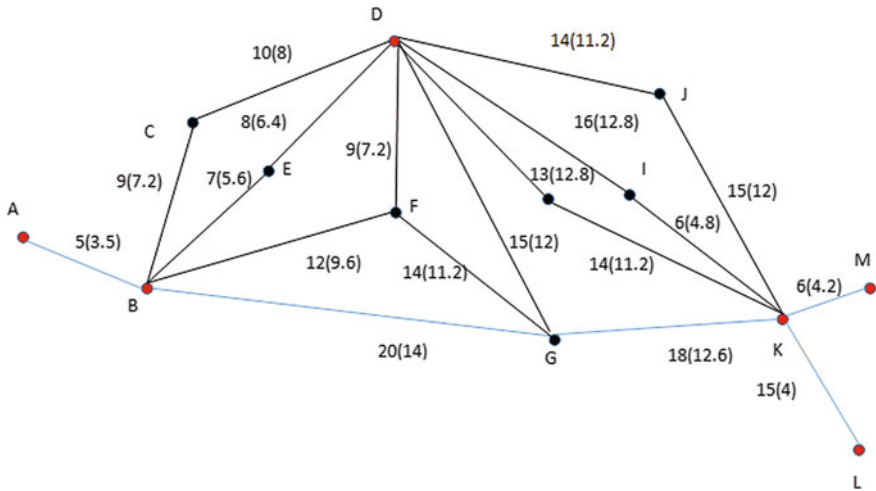
Fig. 1 Partial network of South Chengde road in Huai'an

routes between points B and K. The weights on the route represent the distance and cost between two points. The cost increases by the degree of deviation from the fixed orbit. Length unit is kilometer, and cost unit is yuan.

First of all, we can take the distance, time, and cost as 0.2, 0.3, and 0.5, respectively, according to the questionnaire survey. Then, label the graphs basing on the principle of Dijkstra's algorithm. The label of point B is $P = 0$. T labels are used to mark the other points besides point B. Select the smallest number of all T labels to be changed into P label and cross T label, as shown in Fig. 2.

The shortest path obtained from the graph is $B \rightarrow E \rightarrow D \rightarrow I \rightarrow K$, and the shortest distance is 37. Similarly, the most economical path is $B \rightarrow G \rightarrow D \rightarrow H \rightarrow K$, and the saving cost is 29.6.

Adjacency matrix representation: The image representation of the connection between the nodes is given firstly.

$$D^{(1)} = \begin{bmatrix} 0 & 9 & \infty & 7 & 12 & 20 & \infty & \infty & \infty & \infty \\ 9 & 0 & 10 & \infty & \infty & \infty & \infty & \infty & \infty & \infty \\ \infty & 10 & 0 & 8 & 9 & 15 & 13 & 16 & 14 & \infty \\ 7 & \infty & 8 & 0 & \infty & \infty & \infty & \infty & \infty & \infty \\ 12 & \infty & 9 & \infty & 0 & 14 & \infty & \infty & \infty & \infty \\ 20 & \infty & 15 & \infty & 14 & 0 & \infty & \infty & \infty & 18 \\ \infty & \infty & 13 & \infty & \infty & \infty & 0 & \infty & \infty & 14 \\ \infty & \infty & 16 & \infty & \infty & \infty & \infty & 0 & \infty & 6 \\ \infty & \infty & 14 & \infty & \infty & \infty & \infty & \infty & 0 & 15 \\ \infty & \infty & \infty & \infty & \infty & 18 & 14 & 6 & 15 & 0 \end{bmatrix}$$
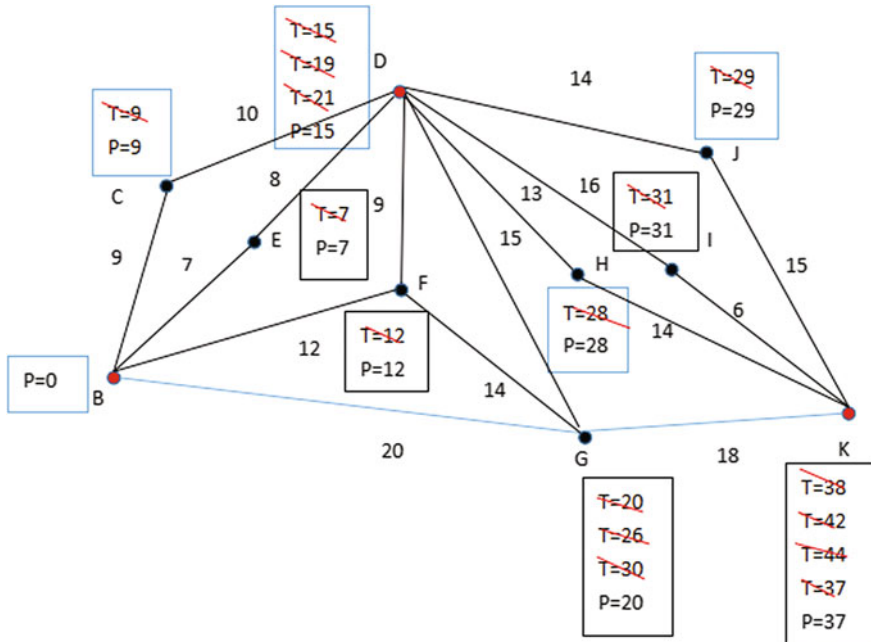
Fig. 2 Short circuit diagrams of the label method

$D^{(2)}$ is obtained through the distance matrix algorithm as follows:

$$
D^{(2)} =
\begin{bmatrix}
0 & 9 & 15 & 7 & 12 & 20 & \infty & \infty & \infty & 38 \\
9 & 0 & 10 & 16 & 19 & 25 & 23 & 26 & 24 & \infty \\
15 & 10 & 0 & 8 & 9 & 15 & 13 & 16 & 14 & 22 \\
7 & 16 & 8 & 0 & 17 & 23 & 21 & 24 & 22 & \infty \\
12 & 19 & 9 & 17 & 0 & 14 & 22 & 25 & 23 & 32 \\
20 & 25 & 15 & 23 & 14 & 0 & 28 & 24 & 29 & 18 \\
\infty & 23 & 13 & 21 & 22 & 28 & 0 & 20 & 27 & 14 \\
\infty & 26 & 16 & 24 & 25 & 24 & 20 & 0 & 21 & 6 \\
\infty & 24 & 14 & 22 & 23 & 29 & 27 & 21 & 0 & 15 \\
35 & \infty & 22 & \infty & 32 & 18 & 14 & 6 & 15 & 0
\end{bmatrix}
$$

In accordance with the algorithm principle, the distance matrix has been calculated until $D^{(k+1)} = D^{(k)}$, that is $D^{(3)} = D^{(4)}$.

$$D^{(3)} = D^{(4)} = \begin{bmatrix} 0 & 9 & 15 & 7 & 12 & 20 & 28 & 31 & 29 & 37 \\ 9 & 0 & 10 & 16 & 19 & 25 & 23 & 26 & 24 & 32 \\ 15 & 10 & 0 & 8 & 9 & 15 & 13 & 16 & 14 & 22 \\ 7 & 16 & 8 & 0 & 17 & 23 & 21 & 24 & 22 & 30 \\ 12 & 19 & 9 & 17 & 0 & 14 & 22 & 25 & 23 & 31 \\ 20 & 25 & 15 & 23 & 14 & 0 & 28 & 24 & 29 & 18 \\ 28 & 23 & 13 & 21 & 22 & 28 & 0 & 20 & 27 & 14 \\ 31 & 26 & 16 & 24 & 25 & 24 & 20 & 0 & 21 & 6 \\ 29 & 24 & 14 & 22 & 23 & 29 & 27 & 21 & 0 & 15 \\ 37 & 32 & 22 & 30 & 31 & 18 & 14 & 6 & 15 & 0 \end{bmatrix}$$

The distance matrix algorithm is used to find the distance between any two nodes. In connected vehicles, the system can run the distance matrix algorithm and the distance between two points can be recorded. Because the real-time location of the customer is on point D, we have to reach point D first. The routes from B to D are $B \to C \to D$, $B \to E \to D$, $B \to F \to D$, $B \to G \to F \to D$, and $B \to G \to D$. And the coincidence routes from point D to K are $D \to J \to K$, $D \to I \to K$, and $D \to H \to K$. The distances on BCD, BED, BFD, $BGFD$, and BGD between B and D are 19, 15, 21, 43, and 35, respectively. The distances on DJK, DIK, and DHK between D and K are 29, 22, and 27, respectively. Similarly, the costs that correspond to BCD, BED, BFD, $BGFD$, BGD, DJK, DIK, and DHK are 15.2, 12, 16.8, 32.4, 26, 23.2, 17.6, and 21.6, respectively.

According to the survey data and historical data, the moment speed of a car on the road can be measured, as well as the traveling speed of free flow, traffic density, and traffic density in a traffic jam, which are 25, 50, 0.35, and 0.7, respectively. According to the formula, $q_k = 8.75$, $t_0 = 0.38$. The time corresponding to BCD, BED, BFD, $BGFD$, BGD, DJK, DIK, and DHK are 26, 24, 27, 35, 28, 29, 30, and 27.

Take the results into the weight calculation:

The normalization of the original data is shown in Table 1.

Table 1 Normalization of the original data

Routes	Distance (d)	Time (t)	Cost (c)
BCD	0.29	0.18	1
BED	0	0	0.33
BFD	0.43	0.27	0.67
BGFD	1	1	1
BGD	0.71	0.36	0.68
DJK	0.93	0.55	0
DIK	0.5	0.36	0.33
DHK	0.71	0.27	0.17

Table 2 Comparison between coordinated route and route without coordination

Carpool routes	Total distance (km)	Total cost (yuan)	Total time (min)
Coordinated route for dynamic uncertain demand	48	37.3	40
Route without coordination for dynamic uncertain demand	78	45.3	51
Saving percentage	38.46	17.66	21.57
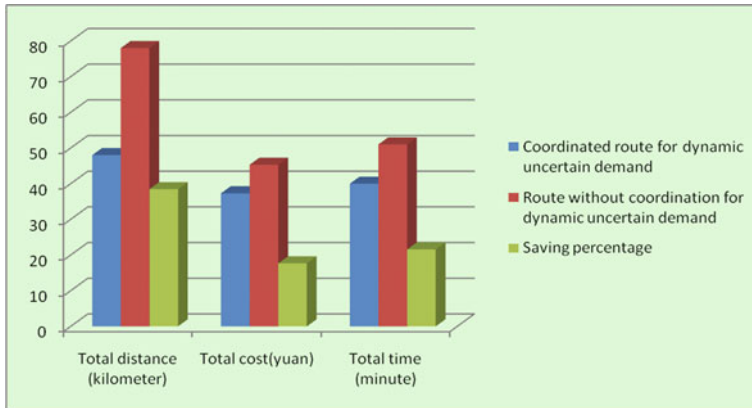
Fig. 3 Contrast on coordinated route and route without coordination

Some data can be obtained according to the formula of weight sample coefficient. The sample weight coefficients that correspond to *BCD*, *BED*, *BFD*, *BGFD*, *BGD*, *DJK*, *DIK*, and *DHK* are 0.612, 0.165, 0.502, 1, 0.590, 0.351, 0.373, and 0.308, respectively. Considering the combined sample weight coefficients, the selected path is $B \rightarrow E \rightarrow D \rightarrow H \rightarrow K$. The shortest path is $B \rightarrow E \rightarrow D \rightarrow I \rightarrow K$. The path with the least time is $B \rightarrow E \rightarrow D \rightarrow H \rightarrow K$. The lowest cost route is $B \rightarrow G \rightarrow D \rightarrow H \rightarrow K$. Finally, these integrated information and all kinds of choices may be feed backed to the customers. Then, the users can choose routes *BED*, *DHK*, *KL* together and also choose the routes with the shortest time, distance, or the lowest cost according to their actual situations. If you choose not to cooperate carpool, then the costs of *KE* and *KL* are paid by the carpool customers. However, compared with the coordinated carpool, it will consume more resources. The specific differences are shown in Table 2 and Fig. 3:

Conclusion

The traditional carpool route selection and optimization methods focusing on single factor have been unable to meet the practical needs. Taking comprehensive consideration on multi-factor will meet the trend of the development of complex

information age. Accordingly, the paper puts forward a method of selecting and optimizing the coordinated carpool route for dynamic and uncertain demand based on connected vehicles. Firstly, calculate the distance, cost, and time of the selected route by Dijkstra's algorithm, then get the distance matrix combined with the cost function and the road resistance function. Following, the concept of sample weight coefficient is introduced. And the models on sample weight calculation and coordinated route selection are built. The results show that comparing with a common route, the selected optimal coordinated carpool route could decrease the total distance by 38.46%, reduce the total cost by 17.66%, and save the total time by 21.57% for dynamic and uncertain demand. Multi-path coordination problem in complex networks using big data mining will be an important direction for the future research.

Conflict of Interests The authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. Xie, Qiuyan. 2012. *Study of Vehicle Dynamic Route Guidance under the Condition of Vehicle Networking*. China: Institutes of Technology of South China.
2. Lin, Xiaoling. 2014. *Research and Implementation of Vehicle Dynamic Shortest Path Based on Genetic Neural Network*. China: University of Fuzhou.
3. Chen, Jingrong., and Ruihua, Xu. 2012. Route Choice Model of Traffic Network with Preference. *Journal of Traffic and Transportation Engineering*.
4. Weng, Kerui., Kejun, Zhu., and Geng, Liu. 2015. *Research on Route Integration of Coordinated Transportation*. China: China Management University, (01).
5. Weng, Kerui., Zi-hao, Xu., and Aderemi Oluyinka, Adewumi. 2014. Flow Merging and Hub Route Optimization in Collaborative Transportation. *Journal of Applied Mathematics*.
6. Oh, Hyondong., Seungkeun, Kim., and Antonyos Tsourdos, Brian A. White. 2014. Coordinated Road- Network Search Route Planning by a Team of UAVs. *International Journal of Systems Science* 45 (5): 825–840.
7. Hawas, Yaser E. 2012. A Cooperative Distributed System for Real-Time Route Guidance. *Journal of Transportation Technologies* 02 (03): 230–240.
8. Mengfei, Wen. 2013. *Research on Key Technologies of collaborative optimization and guidance of traffic flow in Urban Intelligent Transportation System*. China: Central South University, (03).
9. Mateos, Alfonso., Antonio, Jimenez., and Jose.F, Blanco. 2012. Dominance Measuring Method Performance under Incomplete Information about Weights. *Journal of Multi-Criteria Decision Analysis* 19: 3–4.

10. Sun, Zhanquan, Gu Weidong, Yanling Zhao, and Chunmei Wang. 2013. Optimal Path Finding Method Study Based on Stochastic Travel Time. *Journal of Transportation Technologies* 03 (04): 260–265.
11. Zhou, Guiliang, Kai Huang, Lina Mao, and Yanru Zhu. 2014. The Realization and Evaluation of Fixed Time and Fixed Line Carpool to Commute—A Case Study of Huai'an. *Traffic information and security* 4: 41–45.
12. Gang, Li., and Anantaram, Balakrishnan. 2016. Models and algorithms for network reduction. *European Journal of Operational Research* 248(3).

A Research on Traffic Conflict Characteristics of Vehicles Going Straight or Turning Right at Large Intersections

Xinglei Zhang and Xianghai Meng

Abstract Taking the intersection of Zhongshan Road and Zhongxuan Street (the intersection in front of the Municipal Government) as its research object, this paper studies the traffic conflict characteristics of vehicles going straight and turning right at this intersection with post-encroachment time (PET), classifies the conflict characteristics by time and space, and finds out the time periods and space regions for severe converging conflicts with 1.5 s as the threshold of PET.

Keywords Traffic engineering · Large intersection · Traffic conflict · Time characteristic · Space characteristic

Introduction

With the expedition of urbanization, motor vehicles are becoming increasingly popular. People are bothered with traffic congestion and traffic accidents when enjoying the great convenience brought by motor vehicles. As an important part of the micro-level of urban transportation, the function of an intersection is to connect the roads to different directions and form a road network so that vehicles can make free turns and crossings in the network. Traffic flows from different directions converge to the intersection, so it becomes a bottleneck of urban road transportation and a position of frequent traffic accidents. Thus, the effective way to relieve urban traffic congestion and reduce traffic accidents is to implement scientific management to the intersections, fully exert its traffic capacity and ensure the safety of road users.

Traffic conflict technique is a method to analyze and assess traffic safety based on non-accident data, which is different from traditional methods based on the statistical analysis of accident data for its features of large sample, time sensitivity,

X. Zhang (✉) · X. Meng
School of Transportation Science and Engineering, Harbin Institute of Technology, Harbin 150090, Heilongjiang, China
e-mail: zhang_xinglei321@126.com

low social cost, quantitative analysis, and assessment of transportation safety. This technique was originally brought forward by Perkins and Harris from General Motors; it identifies the conflicts according to the phenomena of emergency braking and steering, and the purpose is to assess whether the vehicles produced by GM are safer than those produced by other manufacturers [1].

Data Description

Introduction to the Intersection

This paper focuses on the research on traffic conflict characteristics of vehicles going straight or turning right at large intersections. The representative intersection of Zhongshan Road and Zhongxuan Street (the intersection in front of the Municipal Government) is selected for the research, which is a large cross-shaped intersection with six lanes on both sides, and one of them is for right turn; please refer to Fig. 1 for details. This is a key transportation junction with large traffic flow and serious congestion during morning and evening peak hours. The traffic conflict of vehicles going forward and turning right was not taken into consideration when this intersection was designed. This section was selected for our research because it can reflect the problem of traffic conflict of vehicles going forward and turning right at the section.

Fig. 1 Intersection of Zhongshan Road and Zhongxuan Street (the intersection in front of the municipal government)
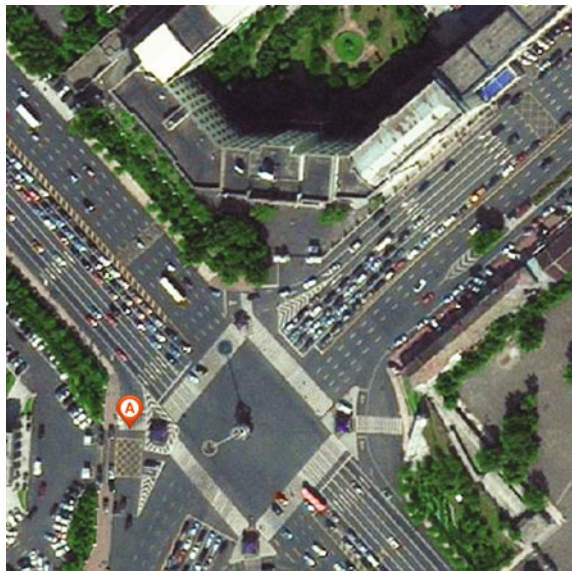
Investigation Method of Conflict

This paper mainly studies the characteristics of post-encroachment time (PET). Field observation is applied, which means the investigators take real-time recording of field traffic conflicts or record traffic conflicts by watching the recorded videos of the field traffic flow. In this paper, the researchers observe traffic conflict by watching the recorded videos and by implementing the data analysis and categorization so as to acquire the data of traffic conflict number, flow rate, and density. Videos recorded during the time period 7:30–16:30 on March 24–26, 2015, are used for this investigation, and all the videos were captured with fine weather and normal traffic.

Analytical Methods of Traffic Conflicts

Conflict Analysis

Traffic conflicts can be categorized in different ways, on the basis of conflict locations, types of road users involved, possible collision modes, and severity of traffic conflicts. Traffic conflicts are likely to occur at intersections and ramps because traffic flows often mingle and intersect there. Besides motor vehicle flows, there are also non-motor vehicle flows and pedestrian flows in the mixed traffic flows, resulting in the four phenomena of crossing, diverting, converging, and confluence, so that conflicts can be categorized into diverting conflict, crossing conflict, converging conflict, and confluence conflict. This paper mainly studies the converging conflict of motor vehicles at intersections.

Analysis of Conflict Severity

Traffic conflicts are usually measured according to approaching degrees of time and space. Two indexes most frequently used for conflict measurement at present stage are time to collision (TTC) and PET.

The precondition for TTC is that conflict process occurs to road users involved, which means if two interactional road users keep their current speeds and directions unchangeable, collision will occur. Due to the difficulty to indentify the conflict process, TTC is more suitable to the situation when the involved road users have the same traveling path, for example, the rear-end conflict (which happens once the speed of the vehicle behind is greater than that of the vehicle in front). The precondition for PET is relatively looser compared with that of TTC. It only requires a conflict point or conflict face between the involved road users, so PET is more frequently used to measure the conflict with intersecting traveling paths,

for example, the conflicts within the range of intersection. Moreover, PET measurement is relatively simple because it does not need speed measurement. So in this paper PET is selected to be the measurement index for traffic conflicts of right turning and straight going.

Analysis of Time Distribution Characteristics

Classification of the Peak and the Flat

In order to study the time distribution characteristics at the intersection, we investigate the related intersection and analyze the time characteristics of its flow. Every 15 min is an interval for calculation and one hour is calculated, so a total of 30 intervals are obtained. Data on flows are transformed into standard units (pcu/h).

The hour with the biggest traffic flow is set to be the peak and the other hours are the flat. The traffic flows in the three days of March 24–26, 2015, is observed. Please refer Table 1 for the traffic flows in detail.

The peak and the flat in each day can be determined by analyzing the above data of traffic flows; please see Fig. 2 for details. It is indicated in the figure that the peak in the three days are the hour from 7:30 to 8:30 and the hour from 16:30 to 17:30, which are also the actual peak hours of mornings and evenings.

Table 1 Traffic flows on the straight-going lane in the three days (pcu/h)

	March 24	March 25	March 26
7:30–8:30	3969	4614	4036
8:30–9:30	3560	4120	3832
9:30–10:30	2971	3223	2684
10:30–11:30	3448	3148	2758
11:30–12:30	3007	3369	3122
12:30–13:30	2858	3337	2937
13:30–14:30	2696	3735	3335
14:30–15:30	2545	2825	2434
15:30–16:30	2758	3575	3198
16:30–17:30	3784	4155	3811

Fig. 2 Traffic flow distribution diagram of the three days
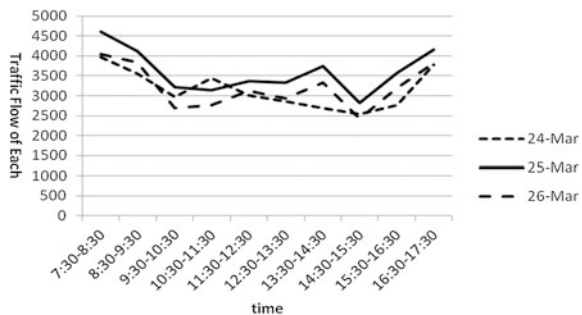
Comparison of Conflict Rates

Conflict rate is introduced as a determining index to analyze the severity more rationally and more reasonably. The calculation formula of conflict rate is as follows:

$$R_{ij} = \frac{T}{\sqrt{P_i \cdot P_j}} \tag{1}$$

In this formula,

R_{ij} The traffic conflict rate between i and j; i and j represent different participants;

T The number of conflicts in each hour at the intersection;

P_i, P_j Equivalent traffic flow of participants i and j

Please see Table 2 for the conflict rates in the three days calculated according to formula (1).

Table 2 indicates that the conflict rates at the morning peak (7:30–8:30) are, respectively, the maximum value for each day, while the conflict rates at the evening peak are not so maximum.

Comparison of Severity Degrees

This paper uses PET as the evaluation index of conflict severity. PET refers to the time difference for two vehicles arriving at the defined section. Taking the converging area at intersections as an example, PET is the conflict time during the converging process of straight going and right turning, and its calculation principle is shown in Fig. 3. As an index to judge the traffic conflict with time as its

Table 2 Conflict rates in the three days

	March 24	March 25	March 26
7:30–8:30	0.027	0.026	0.023
8:30–9:30	0.019	0.012	0.009
9:30–10:30	0.021	0.019	0.019
10:30–11:30	0.023	0.020	0.016
11:30–12:30	0.022	0.008	0.014
12:30–13:30	0.011	0.025	0.015
13:30–14:30	0.022	0.020	0.031
14:30–15:30	0.054	0.006	0.008
15:30–16:30	0.006	0.008	0.017
16:30–17:30	0.016	0.015	0.011
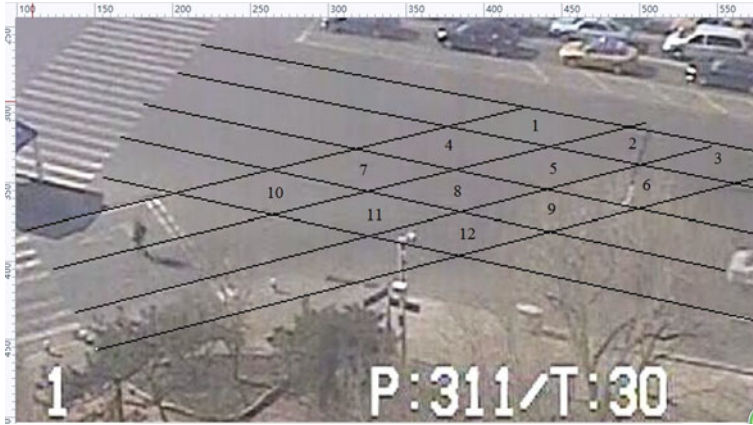
Fig. 3 Grid division at the intersection space

measurement, the application of PET is quite less than that of TTC, and its formulae are as follows:

$$PET_i = t_{i-1} - t_i \tag{2}$$

$$PET_{i-1} = t_i - t_{i-1} \tag{3}$$

In the formulae,

PET$_i$ The time difference (s) for the ith vehicle and the vehicle in front of it, namely the $(i - 1)$th vehicle, to arrive at the defined section;

PET$_{i-1}$ The time difference (s) for the $(i - 1)$th vehicle and the vehicle behind it, namely the ith vehicle, to arrive at the defined section;

t_i The time (s) for the vehicle behind to arrive at the defined section;

t_{i-1} The time (s) for the vehicle in front to arrive at the defined section

On the basis of the PETs for each period in the three days calculated with formulae (1) and (2), and with reference to research achievements at home and abroad, this paper sets 1.5 s as the threshold of PET. When PET is less than 15 s, vehicles will seriously impact each other.

Severity levels of PET at the peak and the flat are compared according to the setting of peak and flat periods. Table 3 shows that the conflict severity during the peak period in the morning (7:30–8:30) is the maximum for each day, while the severity of evening peak period (16:30–17:30) is not so serious.

Table 3 Statistics of peak and flat periods in the three days when PET I less than the threshold

Periods	March 24		March 25		March 26	
	Conflicts number less than 1.5 s	Percentage of total conflict number (%)	Conflicts number less than 1.5 s	Percentage of total conflict number (%)	Conflicts number less than 1.5 s	Percentage of total conflict number (%)
7:30–8:30	55	98.182	30	47.619	48	96.000
8:30–9:30	33	94.286	9	37.500	15	83.333
9:30–10:30	12	38.710	18	56.250	20	71.429
10:30–11:30	30	76.923	13	40.625	11	47.826
11:30–12:30	12	33.333	9	64.286	6	28.571
12:30–13:30	7	41.176	31	70.455	16	66.667
13:30–14:30	11	34.375	19	50.000	30	56.604
14:30–15:30	10	13.333	3	33.333	3	30.000
15:30–16:30	5	55.556	7	50.000	15	51.724
16:30–17:30	30	96.774	15	46.875	18	52.381

Analyses of Space Distribution Characteristics

Region Division

The corresponding intersection is divided into a number of grid units, and traffic conflicts occurring in each unit are assessed so as to confirm the severity. Theoretically, the more specific the grid division is, the more precise the confirmation of road safety situation in each grid unit is. This paper draws the following grid figure according to the number of lanes and vehicle width.

Comparison of Conflict Rates

On the basis of the existing formula of conflict rate, the formula of conflict rate at the intersection of space division is as follows:

$$R_{nij} = \frac{T_n}{\sqrt{P_{ni} \cdot P_{nj}}} \qquad (4)$$

In the formula,

R_{nij} Traffic conflict rate of i and j in region n; i and j represent different participants;

T_n The number of conflicts per hour in region n at the intersection;
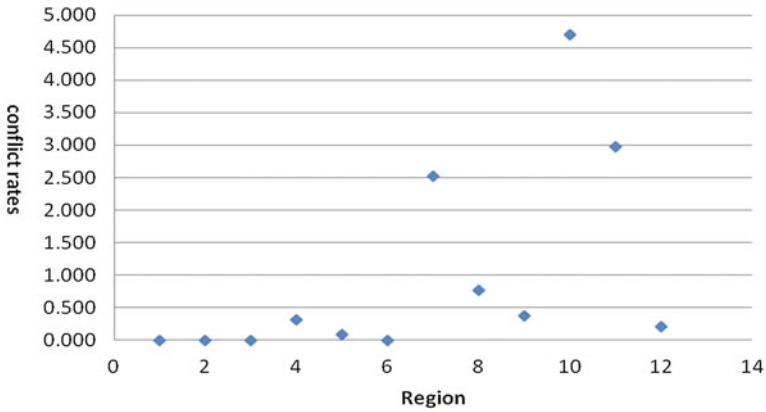
Fig. 4 Conflict rates

P_{ni}, P_{nj} Equivalent traffic volume of i and j in region n;
n   1, 2, …, 12

Please see Fig. 4 for conflict rates at different periods of each region calculated according to formula (4).

Figure 3 indicates that region conflict rate 10 > 11 > 7, 1, 2, 3, and 6 is zero. It shows that conflict of straight and turning right in intersection is mainly concentrated in region 7, 10, and 11.

Comparison of Severity Levels

Please see Table 4 for conflict data with PET less than 1.5 s in each region based on the division of space grid.

Table 4 indicates that conflicts are severe in the regions 7, 10, and 11, and Fig. 3 shows that regions 10 and 11 are the major regions for converging of straight going

Table 4 Distribution of the number of conflicts in different regions

Region number	Number of conflicts		
	March 24	March 25	March 26
4	2	5	2
5	1	4	1
7	47	30	31
8	14	15	13
9	5	2	5
10	80	53	72
11	67	47	60
12	5	2	3

and right turning vehicles at the intersection, while the conflicts in region 7 are severe because some vehicles turning right make a big radius right turn and insert into straight going lane forcibly, so the conflict severity is smaller than that in regions 10 and 11.

Conclusion

With the increase in motor vehicle population, traffic conflicts of vehicles going straight and turning right at large urban intersections are on the rise. It can be seen that the problem at urban intersections focuses on the converging regions of straight going and right turning. This paper analyzes the traffic conflict characteristics of vehicles going forward and turning right at the large intersection and finds out the time periods and regions with severe conflicts so as to provide references for future reconstruction of large urban intersections.

References

1. Perkins, R.S., and J.I. Harris. 1968. Traffic conflict characteristics-accident potential at intersections. *Highway Research Record* 224: 35–43.
2. Li, Yang, Jianxun, Liu, Daji, Shen, Xuan, Liu, and Dong, Qu. 2009. Classification and discretion of urban road traffic conflict. *Highways and Automotive Applications* (6): 50–53.
3. Pei, Yulong, and Shumin Feng. 2007. A research on risk degree of pedestrians crossing street based on traffic conflict. *Journal of Harbin Institute of Technology* 39 (2): 285–287.
4. Xiang, Qiaojun, Chuan Lu, Qun Wu, Jian Lu, and Yongfeng Ma. 2009. A research on the prediction mode of traffic conflicts at non-signal controlled intersection. *Journal of Highway and Transportation Research and Development* 26 (5): 132–135.
5. Zhou, Junchang, Yulin Chang, Min Guo, and Guohua Wang. 2011. Safety assessment of expressways based on traffic conflict technique. *Journal of Chongqing Jiaotong University (Natural Science)* 30 (5): 974–978.
6. Quo, Weiwei, Shaowei Qu, and Dianhai Wang. 2011. Discretion model of traffic conflict. *Journal of Jilin University (Engineering and Technology Edition)* 41 (1): 35–40.

Summary and Development Trend of Traffic Equilibrium Research

Lingmin Yang, Hong Wang and Shunying Zhu

Abstract Traffic congestion is a major problem currently major cities face with, and the root cause of traffic congestion is the imbalance between transportation supply and demand. This article summarizes the cognition and research status of traffic equilibrium, and analyzes the limitations and shortcomings of traffic equilibrium realization mechanism. It points out that the assumptions of existing traffic equilibrium models are too idealistic. These models are not in conformity with the conditions and processes of the equilibrium in reality, and there are drawbacks like not all-sided considering the travelers' features such as individual homogeneity, complete rationality, and perfect information, and models are hard to solve under the "top-down" solution mode, when taking time variation of traffic flow or randomness of travelers' route choice procedure into consideration. It also summarizes the new method of studying the complex adaptive system in social and economic fields and puts forward the idea and prospect that future research on the traffic equilibrium could draw lessons from the related achievements in the researches of social and economic fields. By applying complex adaptive system theory to research of traffic equilibrium and adopting research method of agent-based modeling and simulation under "bottom-up" mode, research in the following future could establish traffic equilibrium model that is in accordance with the characters of traffic participants such as individual heterogeneity, bounded rationality, incomplete information, self-organization, intelligent learning, and decision-making. It opens up a new way to research on traffic equilibrium in the perspective of complex adaptive system theory.

Keywords Traffic congestion · Traffic equilibrium · Traffic equilibrium realization mechanism · Complex adaptive system theory

L. Yang · H. Wang · S. Zhu (✉)
School of Transportation, Wuhan University of Technology, Wuhan 430063, China
e-mail: zhusy2001@163.com

Introduction

Traffic congestion is a major problem currently major cities face with, and the root cause of traffic congestion is the imbalance between transportation supply and demand. Radically speaking, the method to deal with traffic congestion is to coordinate the supply and demand of traffic infrastructure [1], and whether the traffic supply meets with the traffic demand determines whether there is traffic congestion fundamentally.

From the perspective of traffic planning, which viewing from a one-year or several years' period, there is surely an imbalance between the continuous increase of traffic demand and the laddering increase of traffic supply. As shown by the Law of Dams, constructing new roads is not an effective way to deal with traffic congestion, because the newly constructed roads will induce more traffic, and traffic demand volume is always tend to surpass traffic supply eventually [2–5]. This also explains why many cities fall into the vicious cycle of "vehicle number increases—traffic congestion, worse environment—construct roads—vehicle number increases—more traffic congestion, even worse environment" [6]. Generally speaking, it is not applicable to realize the equilibrium between traffic supply and demand and solve the problem of traffic congestion by unilaterally increasing traffic supply.

From the perspective of traffic guidance and control which viewing from a period counted by minutes even seconds, it realizes the relatively balanced traffic flow mode by managing and adjusting the traffic demand under the condition that traffic supply (road resource) is given. As mentioned, realizing the equilibrium of traffic supply and demand by increasing traffic supply is not feasible. Therefore, realizing the equilibrium by adjusting the traffic demand is the most important focus on traffic equilibrium at present, and is the narrowly speaking traffic equilibrium means. This paper discusses the "traffic equilibrium" of this meaning.

The Cognition of Traffic Equilibrium and Its Research Status

The Principles of Traffic Equilibrium: User Equilibrium and System Optimum

The concept of traffic equilibrium starts from the principles of distributing traffic volume with a balanced approach, which was proposed by Wardrop. The traffic equilibrium theory and models developed from it plays a domain role in traffic assignment theory. Under the First Principle of Wardrop (user equilibrium principle), the road travelers chosen take the minimum travel time when traffic reaches equilibrium; under the Second Principle of Wardrop (system optimum principle), total travel time or average travel time of all roads of the studied road network will be minimum. But under the system optimum principle, it is necessary for all

travelers to cooperate to decide the exact route of every participant, so as to reach the goal of system optimum that total travel time or average travel time to be minimum and a particular traveler may lower his travel time by changing his travel path under this circumstance. Thus, cooperation can hardly be achieved, and system optimum cannot describe equilibrium status of travelers' behavior.

The Predicament of Balanced Traffic Assignment Models

In the precondition that road resource is given, methods of distributing the traffic volume onto road network in equilibrium can be divided into two categories: the dynamic traffic assignment (DTA) models and the static traffic assignment models. The static traffic assignment supposes that during the time range the model involving, the demand between each OD (orient and destination) is constant, and its main task is distributing the known constant amount of traffic volume onto the given road network. And in the DTA models, with the traffic volume between each OD changing from time to time, its main task is distributing the changing traffic volume onto the existing road network [7, 8].

From the view of assumptions about travelers' behavior, there are two types of typical static traffic assignment models: user equilibrium (UE) models and stochastic user optimum (SUE) models. UE models are based on the assumption of individual homogeneity (travelers have the same bias and decision-making principles on choosing routes), complete rationality (travelers tend to choose the path of maximum utility and minimum impedance), and perfect information (travelers receive full information about the road network) [9]. SUE models loosen the assumption of perfect information [10] (which means it is impossible for travelers to be fully known of the road traffic impedance before they make their route choices, instead they just estimate it, and it is assumed that the difference between the estimated value and the actual value of the road traffic impedance to be a stochastic variable). The multi-user equilibrium network loosens the assumption of individual homogeneity [11] (different types of travelers' weight between travel time and travel cost differently when choosing their routes). Some researches take uncertainties of transportation system into consideration, such as the uncertainty of road network, deviation of travelers' perception, margin of safety, effective travel time, and buffer time [12–17]. A major defect of static traffic assignment models is that its OD distribution matrix is fixed with time variation, which is acceptable in traffic planning from the prospective of a one-year or several years' period [7, 18], but is not acceptable from the perspective of a period counted by minutes even seconds, because the traffic demand is changing with time mentioning traffic congestion, and the static traffic assignment models can do nothing about describing the time-variant traffic demand.

It is necessary to model dynamic traffic flow for the purpose of traffic guidance and control when traffic congestion occurs, and DTA models become the focus of research because DTA models can better describe the time-variant traffic demand.

Corresponding with the static traffic assignment models, the DTA models can be divided into deterministic dynamic user optimum [19–24] and stochastic dynamic user optimum [25, 26] based on different assumptions of individual homogeneity, complete rationality, and perfect information. It is actually an offline process of distribution of given traffic demand [27], and it embodies the time variation of traffic demand, but not the real-time attribute. If any users do not cooperate, this way of "top-down" mode in obtaining traffic equilibrium could hardly meet the need of real-time traffic guidance and control. Meanwhile, DTA models have the defect of being hard to solve mathematically.

The traffic assignment models lack the cognition of conditions and process of equilibrium realization; thus, the actual traffic flow will not follow the traffic assign principles proposed by the mentioned models. As a result, traffic assignment in equilibrium is just an "ideal" situation which cannot exist in reality.

The Dilemma of Traffic Demand Management in Reality

Since the traffic assignment models are a little too idealistic, can equilibrium be realized by adjusting traffic demand features such as travel routes, time of departure, and trip mode?

Traffic demand management is for the aim of seeking dynamic equilibrium between traffic demand and supply and furthermore solve traffic congestion problem by adjusting traffic demand features such as ravel routes, time of departure, and trip mode [28–31]. However, traffic demand management measures be taken sometimes did not bring expected results, some of them are not accepted by travelers, some even misleading citizens, as pointed out by reference [32]. Other route guidance modes, such as the variable message signs (VMS) which are based on real-time traffic flow and car navigation system applications, have problems such as whether travelers comply with the inducing information and the influence of compliance rate and may lead traffic to a new "best" route and cause new congestion on that route [33].

The reason is that we do not have a clear understanding of the mechanism of whether traffic equilibrium can be realized and how to realize, and under this circumstance, traffic management measures are pointless.

The Status of Research on Realization Mechanism of Traffic Equilibrium

The assumptions of existing traffic equilibrium theory and models are too idealistic to be congruent with realization conditions and process of actual traffic equilibrium. Thus, the theory and models can only limit functioning as reference to the research

on realization mechanism of traffic equilibrium in reality. Traffic demand management measures aimed to realize traffic equilibrium and sometimes cannot bring expected results due to lack of clear understanding of the realization mechanism of traffic equilibrium. In-depth study on figuring out what the external conditions are and how these factors effect on traffic equilibrium realization process can make the formulation and implementation of traffic demand management policy better hit the target and be more objective. This is of great importance not only to the development of traffic equilibrium theory, but also to traffic demand management practice and solving the traffic congestion problem, of course, mechanism of traffic equilibrium realization is the key point of current research on traffic equilibrium.

There have been related studies done from the following aspects.

Dynamic System Models Based on Day-to-Day Dynamics of Traffic Flow

Studies based on day-to-day dynamics of traffic flow mainly focus on the evolution process and mechanism of traffic network from disequilibrium to equilibrium, which reflect the dynamic changing performances of decision-making behavior of traffic choice very well. In recent years, researchers have introduced dynamic system models into researches based on day-to-day dynamics of traffic flow from the perspective of evolutionary analysis. Based on the assumption of individual homogeneity and complete rationality, dynamic system models take into consideration the dynamically changing features such as traffic demand, network volume, travel time, and traffic capacity of actual road network. These models can be divided into two types: deterministic process models and stochastic process models. Generally speaking, the deterministic process models are based on the assumptions of perfect information, that is to say travelers are completely clear about real time situations of the traffic system, in spite of the uncertainties of road network, traffic demand and route choice behaviors. The stochastic process model, on the other hand, considers the randomness of travelers' route choice behaviors and usually depicts the evolution process of traffic system by Markov process. However, it becomes complicated and difficult to solve the stochastic process model mathematically, and it is hard to guarantee the system to be convergent and steady [34–37]. And studies based on the overall traffic flow features of the road network evade analyzing the behavior characters of travelers. As pointed out by reference [38], what related current studies do is just searching a kind of dynamic evolving rules for the traffic system to converge to the Wardrop equilibrium. In the precondition that whether the dynamic process converges or not, or how it converges, there is probability that the actual traffic system may not evolve according to that kind of rules.

The Evolutionary Dynamics Based on Evolutionary Game Theory

Studies based on the theory of evolutionary dynamics, which is part of evolutionary game theory, taking into consideration the individual homogeneity, limited rationality of the same degree, and incomplete information, treating influences and interactions between travelers as relationship of game theory participants, is majorly based on the framework of trade-offs and choices between two participants or a set of one-to-one trade-offs in a group of people [39–43]. These studies regard equilibrium as a result of analysis and self-examination done by travelers based on the pre-conditions that the rules, the level of rationality of participates and payoff function is a common view to everyone. However, it is quite different that the trade-offs and choices between two participants or a set of one-to-one trade-offs from a group of people (travelers) participating in a trade-off simultaneously. The research done by Guan Hongzhi proved that the evolutionary stable state of route choosing trade-offs where a group of travelers participating simultaneously is equivalent to SUE, with travelers having features such as individual homogeneity, limited rationality of the same degree, and incomplete information. However, it assumed that the average utility of each round of trade-offs is common view to each traveler, and all travelers who have the same difference value between his utility and the average utility of the group make their choices in choosing routes with the same probability, which is hardly realized in reality.

The Learning Model Based on Individual Behavior

From the view of learning model based on individual behavior, people can hardly be completely rational in reality and is constantly learning from the environment to improve, make progress, and adapt to the environment. This theory introduces reinforcement learning models, Bayes learning models, and stochastic belief learning models to study the equilibrium realization process based on the traveler's learning behavior on the preconditions such as individual homogeneity, limited rationality, and incomplete information [44–46]. Nakayama set up a time discrete traffic flow evolutionary Bayes learning model, through analysis conducted by stability theory of difference equations, he found out that the overall asymptotic stable point converges to UE in the case of incomplete information, and it is very difficult to converge to UE in the case of perfect information, and draw a conclusion that perfect information may not be beneficial for converging to UE [46]. Liu Tianliang set up a cognition renewal Bayes model considering travelers' risk aversion and applying discrete choice theory to study learning behaviors of travelers and the dynamic evolution rule of traffic flow [47]. Xu Hongli set up a trip decision model based on cumulative prospect theory, taking into consideration constraint time and traveler's risk bias, and discussed the equilibrium outcomes under the stochastic network user equilibrium model [40].

However, the assumption of individual homogeneity of travelers is still different from the real behavior of traffic participants.

The Development and Prospect of Researches on Realization Mechanisms of Traffic Equilibrium

The Development Direction of Researches on Realization Mechanisms of Traffic Equilibrium

On the one hand, the macroscopical status (unimpeded or congested and whether it reaches the equilibrium) of traffic flow is made up by choices of microscopical travelers, and it is necessary to study from the perspective of interactions between the decision-making of traffic participants and the traffic flow status for the understanding about realization mechanisms of traffic equilibrium.

On the other hand, previous study shows that the equilibrium under inadequate consideration of the time-varying and real-time features of traffic flow and the assumption of individual homogeneity, complete rationality, and perfect information is practically meaningless. This demands that the study of realization mechanisms of traffic equilibrium to take into consideration the time-varying and real-time features of traffic flow, the individual heterogeneity, limited rationality and incomplete information, and even traveler's learning, decision making and trade-off process based on his historical experiences and behavioral bias.

Existing research achievements on mechanism of traffic equilibrium realization have recognized the influences and taken into consideration the time-varying and real-time features of traffic flow and the features such as the individual non-homogeneity, not absolute rationalism, and not absolute information, but not yet dealt with these three features simultaneously, although these features do exist simultaneously. The reason is similar to the difficulty to solve situation faced by DTA models in consideration of time-varying or the stochastic process dynamic system models in consideration of randomness, and these complexity factors involving variety, randomness, individuality, and interactions between them are far beyond the ability of description for definitized and simplified models under the framework of classic science. The study of realization mechanisms of traffic equilibrium needs a new breakthrough out of the classic science framework and related research techniques.

Complex Adaptive System Theory and Its Application in Social and Economic Fields

Complex adaptive system theory is a branch of complexity theory, which is called "the science of twenty-first century" by academician Dai Ruwei. Complexity theory

is developing, and other branches such as general system theory, cybernetics, artificial intelligence, synergism, hyper-cycle theory, catastrophe theory, chaos theory, fractal theory, complex network theory, and cellular automaton theory were put forward one after another [48–52]. A scientist from Santa Fe Institute (SFI), Holland, open a new prospect of genetic algorithm and categorizer learning system, and put forward the theory of complex adaptive system (CAS) [53, 54], scientists represented by Gellman put forward the concepts of "agent" and "emergence" [55–57], and Kauffman put forward the concept of "self-organizing" [58, 59], and thus gradually forming a burgeoning and important branch of complexity theory—complex adaptive system (CAS) theory.

CAS theory from the study applied in bioscience and physics, typical complex adaptive systems such as the hive, colony, immune system and the brain of human-beings. Researchers found that these systems form complicated macroscopic phenomena through interactions of simple individuals. Thus, CAS has been recognized to be composed of interactional multiple adaptive individuals (known as agents as well). There is no internal or external leader or controller in a CAS, instead these individuals (or agents) learn and accumulate experiences by interaction between different individual units and changing their behavior according to these experiences; hence, the whole system shows the emergence of a complex but orderly organized status.

CAS theory has been applied widely in social and economic fields such as ASM (Artificial Stock Model) by Arthur and Holland [60], the calculable economy model based on Agent by Lane, the economic system mode (ASPEN) based on Agent by Sandia National Laboratory [61, 62], Agent-based Computational Economics by Tesfatsion, the analogous ASPEN model targeted on economic system and SIMECO model by Zhang [63, 64], the mechanism and model of solving dynamic complexity based on intelligent Agent by Jia [65]. It is worth paying attention that the Los Alamos National Laboratory of the USA is now developing a traffic simulation and analysis system, TRANSIMS, which is supposed to be the application of ideas of CAS into traffic area, and is hoped that it can predict traffic demand by "bottom-up" individual based modeling and simulating by seconds. The development of CAS theory and its fruitful research results in social and economic fields that can be good references in the researches of applying CAS theory to traffic area.

The Prospect of Applying Complex Adaptive System Theory to Research the Realization Mechanism of Traffic Equilibrium

CAS theory adopts agent-based modeling and simulation (ABM) technique [66], which is a "bottom-up" modeling technique, that makes the theory a systematic analysis method which binds tightly the macroscopical phenomenon with microscopical individual behaviors, which observes the macroscopical overall emergence

conducted by agents through certain behavioral rules, and it correlates the travelers' decision-making behaviors with microscope and the traffic flow features with macro scope effectively.

The decision-making under interactions between agents of a CAS is with real time and online. Different agents might have same or different structures, and every agent has its own object and behavioral rules and the ability of decision-making on themselves, and this corresponds to the demand of individual heterogeneity. And limited rationality and incomplete information, learning ability, and adaptability can all be reflected very well by appropriately presuming the rules of getting information and rules of learning and decision-making.

To draw a conclusion, applying Complex Adaptive System theory into modeling route choice traffic system to explore the realization mechanism of traffic equilibrium, with the purpose to find how agents (or travelers) interact with each other through the environment constituted by them, just like the way they interact with each other in real traffic flow. Are there strategies of choice which help to bring about the equilibrium of traffic system and further more relieve or solve the problem of traffic congestion? how to guide travelers follow the strategy, is of great practical significance and prospect.

Conclusion

1. Existing traffic equilibrium theory and models lack the cognition of conditions of equilibrium realization and process, and can only limit functioning as reference to the research on realization mechanism of traffic equilibrium, and transportation demand management measures aimed at realizing traffic equilibrium could not reach the expected goal because of its lack in cognition in traffic equilibrium realization either. The realization mechanism of traffic equilibrium is a key problem that urgently needs to be discovered in the study of traffic equilibrium.
2. Factors of variety, randomness, individuality, such as individual heterogeneity, bounded rationality, incomplete information, self-organization, intelligent learning, and decision-making, in the study of the realization mechanism of traffic equilibrium, are far beyond the description of precise and simplified mathematical model under classic science framework; thus, it is of natural limitation and inappropriate to research on it by models and analysis under classic science framework.
3. Applying CAS theory to researches on traffic equilibrium realization mechanisms could solve the problem of correlating the traveler's decision-making behavior with microscope and the traffic flow features with macroscope, taking into consideration travelers' features such as the time-varying and real-time features of traffic flow, the individual heterogeneity, limited rationality, and

incomplete information, and even traveler's learning, decision-making, and trade-off process based on his historical experiences and behavioral bias, is of great practical significance and prospect.

References

1. Liu, Lan, Weike Lu, and Junsong Ying. 2014. New countermeasures of urban traffic congestion. *Journal of Transportation Engineering and Information* 4 (12): 1–5.
2. Lu, Huapu. 2006. *The theory and method of traffic planning*. Beijing: Tsinghua University Press.
3. Duranton, Gilles and Matthew A. Turner. 2009. The fundamental law of road congestion: Evidence from US cities. NBER Working Paper No. 15376, Issued in September 2009, NBER Program(s): EEE.
4. Downs, Authony. 1992. Stuck in traffic: Coping with peak-hour traffic congestion. Brookings Institution.
5. Downs, Authony. 2003. Still stuck in traffic: Coping with peak-hour traffic Congestion. Brookings Institution.
6. Hu, Xiaojian. 2014. *Traveler's decision making behavior analysis and traffic demand management*, vol. 7. Nanjing: Southeast University Press.
7. Lu, Huapu, Qixin Shi, and Yafeng Yin. 1996. Dynamic traffic assignment model—A review and future. *Journal of Highway and Transportation Research and Development* 6 (13): 34–43.
8. Tao, Chen, and Sengfa Chen. 2004. Dynamic traffic assignment model and its appliance. *Journal of Highway and Transportation Research and Development* 1 (21): 89–92.
9. Lu, Huapu, and Yafeng Yin. 1997. A method of variational inequality for analysing the urban traffic network equilibrium. *Journal of Highway and Transportation Research and Development* 6: 24–29.
10. Daganzo, C.F., and Y. Sheffi. 1977. On stochastic models of traffic assignment. *Transportation Science* 1977 (11): 253–274.
11. Nagurney, A. 2000. A multiclass, multicriteria traffic network equilibrium model. *Mathematical and Computer Modeling* 32: 393–411.
12. Chen, A., and W.W. Recker. 2000. *Considering risk taking behavior in travel time reliability*. Irvine, USA: Institute of Transportation Studies (ITS), University of California.
13. Lo, H.K., and Y.K. Tung. 2003. Network with degradable links: Capacity analysis and design. *Transportation Research Part B* 37 (4): 345–363.
14. Lo, H.K., X.W. Luo, and B.W.Y. Siu. 2006. Degradable transport network: Travel time budget of travelers with heterogeneous risk aversion. *Transportation Research Part B* 40 (9): 792–806.
15. Lam, W.H.K., H. Shao, and A. Sumalee. 2008. Modeling impacts of adverse weather conditions on a road network with uncertainties in demand and supply. *Transportation Research Part B* 42 (10): 890–910.
16. Siu, B.W.Y., and H.K. Lo. 2008. Travel time budget in doubly uncertain transport network: Degradable capacity and stochastic demand. *European Journal of Operational Research* 191 (1): 166–181.
17. Chen, A., and Z. Zhou. 2010. The α-reliable mean-excess traffic equilibrium model with stochastic travel times. *Transportation Research Part B* 44 (4): 493–513.
18. Yang, Dongyuan. 1992. The development and scenario of traffic equilibrium model theory. *China Journal of Highway and Transport* 4: 76–81.

19. Ran, B., R. Hall, and D. Boyce. 1996. A link-based variational inequality model for dynamic departure time/route choice. *Transportation Research Part B* 30 (1): 31–46.
20. Huang, H.J., and W.H.K. Lam. 2002. Modeling and solving the dynamic user equilibrium route and departure time choice problem in network with queues. *Transportation Research Part B* 36 (3): 253–273.
21. Lam, W.H.K., Z.C. Li, H.J. Huang, et al. 2006. Modelling time-dependent travel choice problems in road networks with multiple user classes and multiple parking facilities. *Transportation Research Part B* 40 (5): 368–395.
22. Ren, Hua-ling, and Zi-you Gao. 2007. Dynamic user optimal model with departure time— Application of bilevel programming in dynamic traffic assignment. *Journal of Transportation Systems Engineering and Information* 7 (3): 83–89.
23. Chen, Xing-guang, Jing Zhou, Zhuo-jun Li, and Zhen-tao Zhu. 2009. A multidimensional duo variational inequality model for travel choice. *Journal of Industrial Engineering and Engineering Management* 23 (1): 23–28.
24. Zhang, X.N., and H.M. Zhang. 2010. Simultaneous departure time/route choices in queuing networks and a novel paradox. *Networks and Spatial Economics* 10 (1): 93–112.
25. Lim, Y., and B.G. Heydecker. 2005. Dynamic departure time and stochastic user equilibrium assignment. *Transportation Research Part B* 39: 97–118.
26. Sun, Xiaomei. 2011. Research on dynamic route choice models and methods under multi-sources traffic information, vol. 5. Doctoral Dissertation of Jilin University.
27. Yang, Qinghua, Guoguang He, and Ma. Shoufeng. 2000. Reflection on dynamic traffic assignment. *Systems Engineering* 1: 49–53.
28. Lu, Huapu. 2012. Strategies of supply of urban transport infrastructure and travel demand management. *Urban Transportation. Urban Transport of China* 10(3), May 2012.
29. Huang, Yonggen, and Liege, Hu. 2012. Traffic demand management—Effective ways to realize sustainable development in transportation system. *Systems Engineering* 3.
30. Liu, Zhili. 2006. The necessity and policy of urban traffic demand management. *China Transportation Review* 3.
31. Wang, Feng-yuan, Yin-san Chen, and Nian-xiu Song. 2012. Application of travel demand management in China. *Journal of Traffic and Transportation Engineering* 2(2), June 2012.
32. Li, Zhi-yao, Zhi-cai Juan, and Zong Fang. 2005. Resident travel time choice and congestion pricing policy. *Journal of Traffic and Transportation Engineering* 3(5), February 2005.
33. Liu, Ruilin, Hongzhang Liu, et al. 2016. Balanced traffic routing: Design, implementation, and evaluation. *Ad Hoc Networks* 37: 14–28.
34. Hazelton, M., and D. Watling. 2004. Computation of equilibrium distributions of Markov traffic assignment models. *Transportation Science* 38 (3): 331–342.
35. Davis, G., and N. Nihan. 1993. Large population approximations of a general stochastic traffic assignment model. *Operations Research* 41 (1): 169–178.
36. Hazelton, M., S. Lee, and J. Polak. 1996. Stationary states in stochastic process models of traffic assignment: A Markov Chain Monte Carlo approach. *The 13th Symposium on Transportation and Traffic Theory*. Pergamon: Oxford.
37. Renyong, Guo, and Huang Haijun. 2008. Dynamic evolutionary model of traffic allocation under the situation of. *Journal of Management Science in China* 11 (2): 12–19.
38. Zhang, Bo. 2012. *Research on dynamic route choice behavior based on prospect theory*. Doctoral Dissertation of Shanghai Transportation University. October 2012.
39. Xu, Hong-li, Jing Zhou, and Wei Xu. 2010. Stochastic network user equilibrium and system evolution with dependence on reference point. *Systems Engineering—Theory and Practice* 30 (12): 2283–2289.
40. Qi, Hang, Shoufeng Ma, et al. 2015. Experiments on individual strategy updating in iterated snow drift game under random rematching. *Journal of Theoretical Biology* 368: 1–12.
41. Hollander, Y., and J.N. Prashker. 2006. The applicability of non-cooperative game theory in transport analysis. *Transportation* 33 (5): 483–496.
42. Yang, F. 2005. *An evolutionary game theory approach to the day-to-day traffic dynamics*. Madison: University of Wisconsin-Madison.

43. Li, Zhenglong. 2003. A study of route choice behavior of drivers based on the evolutionary game under the condition of traffic flow guidance. *Journal of Transportation Systems Engineering and Information* 3 (2): 23–27.
44. Avineri, E., and J.N. Prashker. 2006. The impact of travel time information on travelers' learning under uncertainty. *Transportation* 33 (4): 393–408.
45. Ben-Elia, E., I. Erev, and Y. Shiftan. 2008. The combined effect of information and experience on drivers' route-choice behavior. *Transportation* 35 (1): 165–177.
46. Nakayama, S. 2009. Bayesian learning, day-to-day adjustment process, and stability of Wardrop equilibrium. *Transportation and Traffic Theory* 21: 425–440.
47. Liu, Tian-liang, Hai-jun Huang, and Jian Chen. 2008. Evolution of day-to-day route choice behavior considering risk aversion and perception updating. *Journal of Traffic and Transportation Engineering* 8 (4): 90–94.
48. Mitchell, Melanie. 2011. *Complexity*, vol. 8. Hunan: Hunan Science and Technology Press.
49. Holland, John. 2000. *Hidden order—How adaptation builds complexity*. Shanghai: Shanghai Science and Education Press.
50. Morin, Edgar. 2001. *Science avec conscience: LIbrairie Ariheme Fayard*. Beijing: Beijing University Press.
51. Auyang, Sunny Y. 2002. *Foundations of complex system theories*, vol. 10. Shanghai: Shanghai Science and Technology and Education Press.
52. Jin, Wu-lun and Yuan-lin Guo. 2004. The sciences of complexity and their evolution 1(1).
53. Holland, J.H. 1986. A mathematical framework for studying learning in classifier systems. *Physica D: Nonlinear Phenomena* 2 (1–3): 307–317.
54. Holland, J.H. 1992. *Adaptation in natural and artificial systems: An introductory analysis with applications to biology, control, and artificial intelligence*. Cambridge, MA: MIT Press.
55. Arthur, B.W. 1996. Increasing returns and the new world of business. *Harvard Business Review* 74 (4): 100–109.
56. Gell-Mann, M. 2006. Consciousness, reduction, and emergence. *Annals of the New York Academy of Sciences* 929 (1): 41–49.
57. Cowan, G.A., D. Pines, and D. Meltzer. 1999. *Complexity: Metaphors, models, and reality*. Jackson, TN: American Perseus Books.
58. Bak, P., C. Tang, and K. Wiesenfeld. 1987. Self-organized criticality: An explanation of 1/f noise. *Physical Review Letters* 59 (4): 381–384.
59. Kauffman, S. 1993. *The origins of order: Self organization and selection in evolution*. Oxford: Oxford University Press.
60. Arthur, W., J. Holland, B. Lebaron, et al. 1997. *Asset pricing under endogenous expectations in an artificial stock market. The economy as an evolving complex systems*, 15–44. New York: Westview Press.
61. Lane, D.A. 1993. Artificial worlds and economics, part I. *Journal of Evolutionary Economics* 3 (2): 89–107.
62. Basu, N., R.J. Pryor, T. Quint, et al. 1998. Aspen: A microsimulation model of the economy. *Computational Economics* 12: 223–241.
63. Zhang, Shi-wei, and Dong-kui Zhao. 2005. A subject-based macro economy model. *Journal of Management Sciences in China* 8 (2): 7–12.
64. Zhang, Jiang, and Xue-wei Li. 2005. Artificial economy model: Modeling complex economic systems. *Complex Systems and Complexity Science* 2 (4): 37–45.
65. Shao, Yan-hua, and Ming-sheng Zhang. 2012. Research on a class of complex adaptive system modeling. *Computer Engineering* 38 (1): 253–255.
66. Miller, John H., and Scott E. Page. 2012. *Complex adaptive systems: An introduction to computational models of social life*. Shanghai: Shanghai People Press.

Traffic Network Structure of Internet of Vehicles

Linghui Xu, Jia Lu and Jian Zhang

Abstract Internet of Vehicles (IoV) is an extended application of Internet of Things (IoT) in transportation. The traffic network is composed of vehicles, roads, persons, and other entities, with wireless interactions between them. In order to get acquainted and learn more about its components and information transfer, this paper puts forward a general Internet framework of IoV. In addition, communication technologies are compared and discussed in detail. Besides, data transmission of each link in the network is analyzed briefly. Finally, remarks of technical challenges are provided, to have a deeper understanding of the status of IoV.

Keywords Traffic problems · Network framework · Communication interaction · Technical challenge

Introduction

Over these decades, there has been an intensified conflict between the rapidly increasing car ownership and the sluggish development of transport facilities. As a result, traffic problems get worse to hold back the development of urbanization.

L. Xu (✉) · J. Zhang
Jiangsu Key Laboratory of Urban ITS, Southeast University, Nanjing 210096, Jiangsu, China
e-mail: xu_linghuibest@163.com

L. Xu · J. Zhang
Jiangsu Province Collaborative Innovation Center of Modern Urban Traffic Technologies, Southeast University, Nanjing 210096, Jiangsu, China

L. Xu · J. Zhang
Jiangsu Province Collaborative Innovation Center for Technology and Application of Internet of Things, Southeast University, Nanjing 210096, Jiangsu, China

L. Xu · J. Zhang
Research Center for Internet of Mobility, Southeast University, Nanjing 210096, Jiangsu, China

J. Lu
School of Transportation, Southeast University, Nanjing 210096, Jiangsu, China

Traffic problems include traffic jams, collision accidents, energy consumption, and air pollution.

Actually, with advanced wireless communication technologies, these traffic issues are likely to be dealt with in a connected environment. In this connected network, each vehicle is assumed to be "visual" and controllable via wireless communications with other objects. According to above assumption, the concept of IoV is gradually proposed and has been attracting more and more attention.

The idea of developing wireless communications in transportation system has captured much attention from researchers since the 1980s. This development was contributed by several factors, including the wide-adopted and relatively low-cost wireless LAN technologies, the application of wireless communications to reduce the safety and environment problems caused by traffic, and the support from various groups and organizations to encourage the development of automatic vehicles and connected vehicles equipped by wireless devices. However, the existing communication network is not suitable for the direct interaction between vehicle and vehicle or between vehicle and infrastructure [1].

Vehicular ad hoc network (VANET) is subsequently raised to integrate the capabilities of new-generation wireless networks to vehicles [2]. It offers direct communication between vehicles and roadside units (RSUs). Also, it makes it possible for traffic control centers and vehicles to spread and achieve real-time traffic conditions within certain scope.

Nevertheless, conventional VANET just includes a limited number of connected vehicles in the mobile network. It makes no sense to collect state information of these samples as the distribution is stochastic and irregular. In special cases, equipped vehicles are even distracted by traditional vehicles in this network. In contrary, IoV develops a complete communication network in a wider area [3]. Under this circumstance, various traffic elements such as drivers, vehicles, and infrastructure are connected via wireless communication technologies on the road.

In fact, the concept of IoV was firstly proposed by the vision of ITS Strategic Research Plan (2010–2014). In particular, it is intended to develop a comprehensive, intelligent traffic communication system in certain countries, which combines different kinds of cars, people, and infrastructure. In another word, it aims to build a connected vehicle environment for the nation [4].

For the sake of establishing the environment of IoV, several countries have developed communication systems one by another. The Intelligent Vehicle-Highway System (IVHS) in the USA integrates roads, persons, and vehicles closely, to significantly increase transportation safety, system efficiency, environment quality, and energy utilization [5]. Apart from that, Japan opened the Vehicle Information and Communication System (VICS) to collect real-time traffic information from local police and road management departments, including traffic jams, routes, parking space, and road accidents. And then, the proposed information is transmitted to the passing vehicles via radio devices [6]. Besides, the telematics has been comprehensively developed and applied all over Europe to build a transport private wireless communication network. It makes information transfer come true, between vehicles, drivers, passengers, service providers, and so forth [7].

Obviously, the development of IoV has been an inevitable trend in transportation. With wireless communication links in the network, state information of each component can be collected and processed in real time. Based on computation results, notifications, warnings, or even mandatory controls will be broadcast to connected vehicles. As a result, a coordinated and sustainable transportation system is built, implementing cooperation and interaction between people, vehicles, roads, and the external environment.

The remainder of the paper is structured as follows. Section "Network Model" proposes a general network model of IoV to comb roughly through network components and communication connectivity. Section "Communication Technology" analyzes and compares several common wireless communication technologies. Section "Component Connectivity" introduces connectivity between vehicles and other things in detail. Section "Challenges" describes the general technical challenges to be solved in the development of IoV. Section "Conclusion" closes the paper with a conclusion and further prospects.

Network Model

As shown in Fig. 1, a traffic network structure of IoV is presented, taking an intersection for example. The traffic network is composed of interested entities and connectivity between them. The entities contain vehicles, infrastructure (RSUs), people, and onboard devices. These objects are linked with each other by
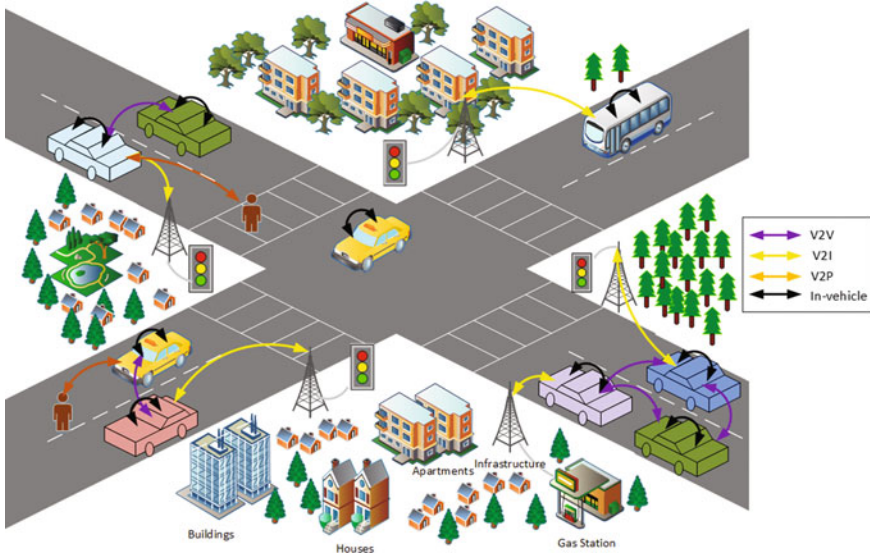
Fig. 1 Network model of Internet of Vehicles

cooperative communications such as Vehicle-to-Vehicle (V2V), Vehicle-to-Infrastructure (V2I), Vehicle-to-Person (V2P), and Vehicle-to-Sensors (V2S) [8].

Vehicles in this network are considered to be all cars that are involved in the IoV. These vehicles are all embedded with onboard unit (OBU). It is the basis for communication between each vehicle and other mobile devices.

RSUs communicate with mobile devices via wireless communications. It collects and stores data from surrounding public facilities, as well as nearby vehicles, at regular intervals. In specific applications, the feedback from RSUs to vehicles is dependent on computation results from a control center, which is connected to RSUs [9].

RSU mentioned above is the core equipment of roadway infrastructure. Infrastructure in the network collects different messages from humans, vehicles, and traffic control centers. All of this information is analyzed and processed to optimize the control strategies or improve characteristics of vehicles.

Persons in traffic network refer to all of the individuals who have access to or have impacts on IoV. It means that persons such as drivers, passengers, and pedestrians are all taken into consideration.

Sensors in the environment are composed of sensing elements installed in vehicles and road detectors. Sensing elements in vehicles gather vehicle attribute values in real time. For example, vehicles' instantaneous velocity can be detected and transmitted via speed detector, generally installed in the transmission case. Apart from that, location data of each vehicle are collected by GPS receiver in vehicles. In an IoV environment, traffic conditions are mostly formed based on data from connected vehicles and other objects. Therefore, the number of road detectors needed is so small that external sensors are ignored in the following discussion of the paper.

Communication Technology

The utilization of wireless communication technologies makes Internet of Vehicles system different from conventional intelligent transportation system (ITS). On the basis of comprehensive comparisons from Crash Avoidance Metrics Partnership, and Dar et al. [10, 11], the capabilities of some common wireless technologies are demonstrated in Table 1.

Table 1 Comparison of common wireless technologies

Communication capabilities	Communication technologies			
	2.5G/3G	GPRS	Bluetooth	DSRC
Directionality	2-way	2-way	2-way	2-way
Latency	~ 110 ms	1.5–3.5 s	~ 100 ms	200 µs
Data rate	1–32 Mb/s	80–384 kb/s	1 Mb/s	~ 6 Mb/s
Range	15 km	10 km	~ 10 m	~ 1 km

Among these communication capabilities in Table 1, communication range and latency are always taken into account when choosing communication technologies to satisfy the requirements of traffic system. According to the above two operating characteristics, four communication technologies are analyzed, including bluetooth communications, satellite communications, 2.5- and third-generation cellular technologies, and dedicated short-range communications (DSRC) [9].

With bluetooth communications, a wireless technology, data transmission can take place between main device and other devices at any time. Therefore, it is likely for objects in transportation system to have a two-way conversation. Bluetooth has been popular for many years and applied in various fields. Its communication latency is a little higher, around 100 ms, which is only used for the interaction between two stable elements.

Comparatively, the scope that GPRS communication can cover is much wider than some other technologies. So it is widely considered that GPRS technology can provide two-way communication between vehicle and vehicle or roadside infrastructure. However, there is only small amounts of information that can be transmitted by this communication. In addition, its communication latency is so high that it is not suitable for real-time utilizations with security purposes.

The 2.5G and 3G communications are the most prevalent among cellular technologies. For example, it is widely used in the information sharing between mobile phones. Nevertheless, connections are likely to be delayed for more time than others, which really leads to distractions. As the 2.5G or 3G technology has a greater latency, it is usually utilized for offline analysis of traffic conditions. Also, it cannot be applied for the real-time traffic safety.

The main benefit of DSRC is that it is applicable to exchanging of plenty of data from transportation system. As a result, information collected form vehicles, humans, and control centers can be thoroughly relayed and discussed. Moreover, this communication has the lowest communication latency, which helps to reduce delays and errors. Besides, its range in a transportation network is perfect for the direct communication among traffic elements.

In a comparison of these four communication technologies, the advantages of DSRC are the most evident in applications of IoV. It has attracted widespread concern and had mature applications in automatic vehicle identification (AVI) and electronic toll collection (ETC). In the rest of this paper, DSRC is assumed to be the only communication technology applied in IoV, to simplify the description of communication connectivity.

SAE J2735 standard describes a message set related to DSRC. In general, the data types and contents of messages in this standard are utilized when dedicated short-range communications is applied for a connected vehicle environment. Although this standard concentrates on DSRC, it is also compatible with other communication technologies. In the newest standard modified in 2016, there are 17 messages, some of which will be mentioned in the following component connectivity part.

Component Connectivity

Vehicle-to-Everything (V2X) is the key technology of IoV. Via communication links, vehicle status is transmitted to other test devices in real time. At the same time, information on traffic conditions and a stream of traffic management and control are sent to vehicles as soon as possible. Based on these two-way interactions, it is capable for vehicles and other components in the network to coordinate with each other.

Connectivity Between Vehicle and Vehicle

Vehicle-to-Vehicle (V2V) communication exists between connected vehicles. As cars come into the network, they become nodes that communicate with others that are close in proximity to them. That is to say, each vehicle with OBU automatically shares its status to other equipped ones in a high frequency, like 10 times one second. The messages sent, through wireless networks, include vehicles' speed, direction, position, routes, and other information at that moment. Consequently, drivers/vehicles are informed of the approaching threats and instructions of braking or steering to avoid collisions. As well, a faster way can be found to reduce travel times, based on surrounding traffic distribution conditions.

V2V communication has been mainly applied to safety improvement and mobility promotion in existing studies, like anti-collision warning application [12] and cooperative automotive platooning application, respectively [13]. Here, only information transfer in the first scenario is analyzed in the following part due to limited space.

As shown in Fig. 2, the incident vehicle shares alert information which provides the details of the incident such as position and moment. This useful information is then conveyed to other equipped vehicles under a certain scope. As a result, more upstream vehicles farther along the road will catch this warning message via V2V communication. The warning message generally fades as the incident has been relieved and handled. Sometimes, this message is closed by an event handler when the incident has no impact on other vehicles. In this scenario, Basic Safety Message
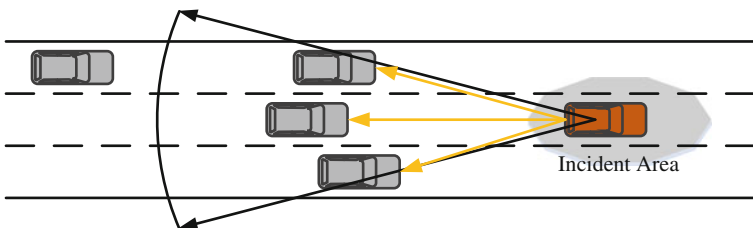
Fig. 2 Communication between vehicle and vehicle in incident warning application

(BSM) in the SAE J2735 standard is required for this application. The BSM mostly contains position information of the incident vehicle as well as its real-time motion state.

Connectivity Between Vehicle and Infrastructure

V2I communication ensures that vehicles and roadside infrastructure are able to communicate with each other. Firstly, messages about real-time traffic conditions are drawn by infrastructure. Some traffic control strategies like intersection signal control steps and some emergent situations like accidents are wirelessly collected by it. After this information is processed, the results are then disseminated to vehicles to remind them of the real traffic around them.

This communication technology has been used for collision avoidance application for safety reasons, like collision avoidance system [14]. Also, it is widely applied in mobility applications to increase vehicle running speed and improve traffic flow. Accordingly, several concepts have been developed, such as signal timing optimization [15], cooperative adaptive cruise control [16], and freeway merge control [17]. For limited space, only the scenario of bus signal priority with information transmission is discussed in detail [18].

As is shown in Fig. 3, the approaching bus transmits its status (e.g., speed, location, number of passengers, and bus ID) wirelessly to an RSU for decisions of signal priority. The RSU is connected to a regional Traffic Management Center (TMC), which takes real-time control over signal timings and phases. The signal priority requests from public transit expire after either it catches the feedback from the RSU or it crosses the stop line.

So as to support this process, one of the most important messages in SAE J2735 standard is the Signal Phase and Timing (SPaT) message. It is utilized to transmit the data of signal phases and timing to other devices. In this case, given the Map Data (MAP) message, the communication devices in vehicles can learn the real signal condition in terms of movements and lanes. Also, the Signal Request
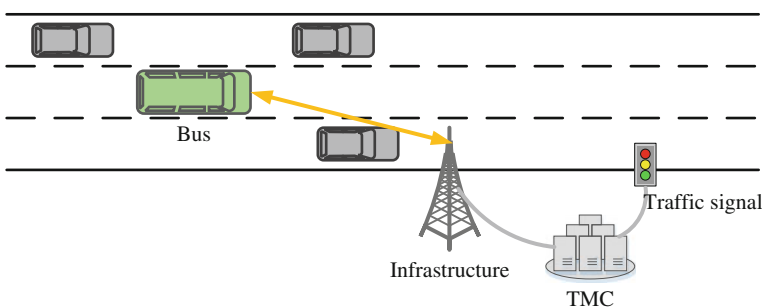
Fig. 3 Connectivity between vehicle and infrastructure in bus signal priority scenario

Message (SRM) is usually sent by certain vehicles to apply for the signal priority. So the SRM is mainly made up of the signal request and the state information of the target vehicles. After the signal request message is completely processed, the intersection control center usually responds to the vehicle in the way of the Signal Status Message (SSM) especially when several request messages are incompatible. However, if there is no SSM, which means limited signal priority request, control strategies with signal phases and timing will be relayed to specific vehicles in the form of SPaT [9].

Connectivity Between Vehicle and Person

Persons in IoV take several forms, including drivers, passengers, pedestrians, and bicyclists. However, V2P communication here only takes walkers and bicyclists into consideration, both of which are collectively called pedestrians. In the traffic network, pedestrians with wireless mobile device exchange their status with vehicles equipped with OBU in a limited range. In another word, both pedestrians and vehicles are capable of sensing potential threats from vehicles/persons in the surrounding environment. Consequently, the number of car-pedestrian collision accidents at blind area will be significantly reduced. Apart from that, wireless communications between vehicles and waiting men make target car "visual" in traffic flow, within a certain range. Thus, a great deal of recognition and waiting time is saved.

With the growing number of casualties and property losses caused by vehicle-to-pedestrian collision accidents, this technology is attracting much intention from researchers. Apparently, the V2P communication is mainly used for road safety, as is expressed by Sungwon Lee et al. [19], Jose Javier Anaya et al. [20], and Lorenzo Caminiti et al. [21], respectively.

To vividly describe a scenario of safety applications, message exchange between vehicles and pedestrians in conflict area is represented in Fig. 4.
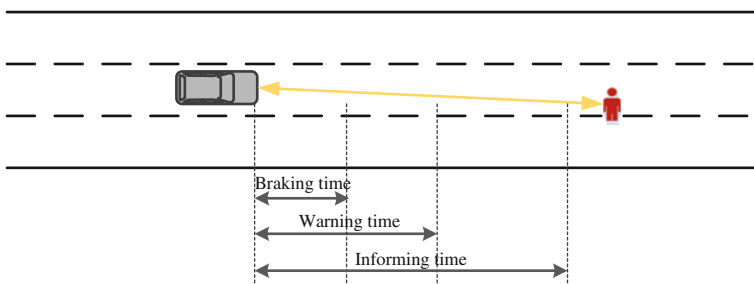
Fig. 4 Communication between vehicles and pedestrians in conflict area

Referring to above schematic illustration, a vehicle travels along a roadway. If a pedestrian is crossing this road within the range, its position information will be sent to the OBU in the vehicle. In an emergency, an alert is going to be generated based on the distance between the pedestrian and the vehicle. The alert to the vehicle varies in degree in three different zones, as shown, informing zone, warning zone, and braking zone [22]. Meanwhile, the pedestrian is reminded of the approaching vehicle. He/she can remain or change his/her speed according to the actual situation. The alert transmission expires after either the person gets over the line of the vehicle or stops at an anti-collision area.

Messages in the SAE J2735 standard support this application as well. Similar to V2V communication, the BSM is composed of different types of vehicle state information such as the locations of vehicles and pedestrians, the motion states of them, and even some warning messages when in an emergent situation.

In-Vehicle Communication

In-vehicle communication is realized by OBU installed in vehicles. It refers to the connectivity between OBU and in-vehicle sensors, as well as internal service devices.

As shown in Fig. 5, OBU basically includes a control unit, a two-way wireless communication system, and a human–machine interface section. The two-way wireless communication system helps the control unit receive and send various signals to other equipped components and systems near this vehicle. The control unit is intended to receive detection signals from various in-vehicle sensors including, but not limited to, an ignition switch sensor, a vehicle speed sensor, and an acceleration sensor. Moreover, the control unit, receiving position information from GPS receiver, is connected with the human–machine interface section to carry
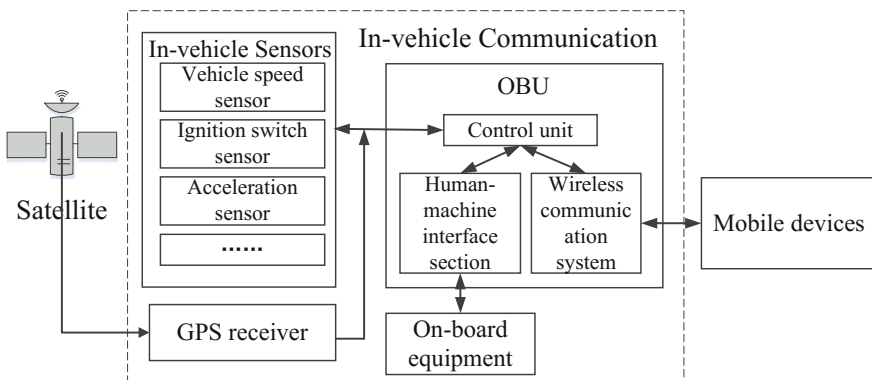
Fig. 5 In-vehicle communication diagram

out navigation controls as well as the implementation of various safety measures and recreational activities [23].

In-vehicle communication is frequently implemented whenever vehicle state information or onboard equipment is needed. In general, diverse sensors, coupled with GPS receiver, send messages to OBU, in which BSM is formed to be transmitted to other surrounding DSRC devices. And when additional information is required via Common Safety Request (CSR) message by another mobile device, the communication becomes two ways between sensors and OBU. In more detail, OBU asks sensors for extra information and receives it in real time.

Challenges

The objective of IoV is to provide connected communication capability to vehicles, humans, and other things to ensure the traffic network manageable, controllable, and credible. Accordingly, a system integrating and coordinating the operations of all parts is expected to be built, to guarantee a safe, efficient, and sustainable connected environment.

However, it will be a long journey before the completion of such a network system. As the Volpe Center estimated in 2008, only about 50% vehicles on the road would be equipped in the following 9 years, no matter what mandatory measures the federal government had taken to encourage the installation of DSRC devices in new cars [24].

Except for low penetration rate of connected vehicles, some technical challenges are discouraging the process of integration. Firstly, the stability and reliability of communication technology are urgent to be improved. The disordered work of delivery system in any link may easily destroy the whole balance and make the manipulation of the network in chaos. Moreover, poor quality of data, as well as transmission latency, is another constraint. An amount of undesirable data with low resolution and errors leads to a large amount of processing time in collection centers. As a result, warnings or control measures are relayed to vehicles with high latency, which may cause collisions. Thirdly, the standards of DSRC are urgently unified. The fact that various DSRC standards exist in different countries is inimical to the long-term and stable development of IoV as standard specifications make it come true that onboard communication equipment made by various manufacturers interacts with each other in most cases. Besides, the privacy requirements of a private vehicle and human being are a critical issue, needed to be settled before large-scale application of connected network. The public may take drastic action against being monitored with destroyed privacy. To the horror of them, their privacy may be pinched by criminals to make them and their family members in danger. The traffic network system may also be invaded by terrorists. If the transmitted data or the control instructions are maliciously tampered, consequences would be devastating.

Therefore, significant breakthroughs in wireless communication technology are crucial to the development of IoV. These breakthroughs are to be realized with mutual efforts and cooperation of government, researchers, manufacturers, and the public.

Conclusion

This paper presents a brief network structure of IoV. The network model is made up of vehicles, humans, infrastructure, and wireless communications among them. Wireless communication technology is the most important element for building a connected environment of IoV. Bluetooth communications, satellite communications, cellular technologies, and DSRC are in common use, of which the first three ones are limited for range or latency issues. Comparatively, DSRC has a wider foreground of applications. Wireless communication connectivity between network components is mainly presented in the form of V2V, V2I, V2P, and in-vehicle communication. Several scenarios and messages are discussed in detail corresponding to these connectivity forms. However, challenges in the low penetration rate of connected cars are leading to the slowing development of IoV. In addition, technical deficiencies of instability, error data, mixed standards, and privacy risks are serious impediments to the progress of transportation.

In summary, transportation has entered the new era of historical development, Internet of Vehicles being the inevitable trend. For the sake of a manageable, controllable, and credible traffic network system, driving forces from various directions are needed. In particular, efforts from researchers, governments, and manufacturers (e.g., theoretical research and applied science, encouraging policies of connected vehicle, and standardization of automobile production) are urgently required.

References

1. Dorrell, D., A. Vinel, and D. Cao. 2015. Connected vehicles-advancements in vehicular technologies and informatics. *IEEE Transactions on Industrial Electronics* 62 (12): 7824–7826.
2. Li, F., and Y. Wang. 2007. Routing in vehicular ad hoc networks: A survey. *IEEE Vehicular Technology Magazine* 2 (2): 12–22.
3. Yang, F., and S. Wang, et al. 2014. An overview of internet of vehicles. *Communications China* 10: 1–15.

4. Grace, N., C. Oxley, and S. Sloan, et al. 2012. *Transforming Transportation through Connectivity: ITS Strategic Research Plan, 2010–2014 (progress update, 2012)*. Its Program Applications.
5. Smith, B.L. 1993. Intelligent vehicle-highway system (IVHS) activities in the Virginia Department of Transportation. *Functional Analysis & Its Applications* 46 (2): 121–132.
6. Fujita, O. 2010. *Vehicle Information and Communication System*. Betascript Publishing.
7. Hübner, U. 2010. *Health telematics Europe. Nursing informatics*, 375–400. London: Springer.
8. Miller, J. 2008. Vehicle-to-vehicle-to-infrastructure (V2V2I) Intelligent Transportation System Architecture. In *Intelligent Vehicles Symposium, 2008 IEEE*. IEEE: 715–720.
9. Zeng, X., K. Balke., and P. Songchitruksa. 2012. *Potential Connected Vehicle Applications to Enhance Mobility, Safety, and Environmental Security*. Southwest Region University Transportation Center.
10. Crash Avoidance Metrics Partnership. 2005. *Vehicle Safety Communications Project Task 3: Final Report-Identify Intelligent Vehicle Safety Applications Enabled by DSRC*. DOTHS-809-859. CAMP Vehicle Safety Communications Consortium.
11. Dar, K., M. Bakhouya, J. Gaber, et al. 2010. Wireless communication technologies for ITS applications [topics in automotive networking]. *Communications Magazine IEEE* 48 (5): 156–162.
12. Cho, W., K.S. Han., and H.K. Choi, et al. 2009. Realization of Anti-collision Warning Application using V2V Communication. *Vehicular Networking Conference (VNC), 2009 IEEE*. IEEE: 1–5.
13. Bergenhem, C., R. Johansson., and E. Coelingh. 2014. V2V communication Quality: Measurements in a Cooperative Automotive Platooning Application. *SAE World Congress 2014*.
14. Federal Highway Administration. 2010. Cooperative Intersection Collision Avoidance System-Stop Sign Assist: Concept of Operations. http://www.dot.state.mn.us/guidestar/ 2006_2010/cicas/CICAS_SSA_ConOps_FINAL_3_18_08.pdf. Accessed Nov 2010.
15. Andrews, S., and M. Cops. 2009. *Vehicle Infrastructure Integration Proof of Concept— Infrastructure*. FHWA-JPO-09-038, FHWA-JPO-09-048, FHWA-JPO-09-057. Washington, D.C.: US DOT Research and Innovative Technology Administration.
16. Federal Highway Administration. 2011. Increasing Highway Throughput Communications and Control Technologies to Improve Traffic Flow. http://www.fhwa.dot.gov/advanced research/pubs/inchwyfact.cfm. Accessed Jan 2011.
17. Smith, B., and B. Park. 2011. Advanced Freeway Merge Assistance: Harnessing the Potential of Intellidrive. http://cts.virginia.edu/research_view_detail.php?projectID=166. Accessed Jan 2011.
18. Liao, C.F., G.A. Davis, et al. 2007. Simulation study of a bus signal priority strategy based on GPS/AVL and wireless communications. *Transportation Research Record: Journal of the Transportation Research Board* 2034: 82–91.
19. Lee, S., and D. Kim. 2016. An energy efficient vehicle to pedestrian communication method for safety applications. *Wireless Personal Communications* 86 (4): 1845–1856.
20. Anaya, J.J., P. Merdrignac., and O. Shagdar, et al. 2014. Vehicle to Pedestrian Communications for Protection of Vulnerable Road Users. *2014 IEEE Intelligent Vehicles Symposium (IV)*, 1037–1042.
21. Caminiti, L., J.C. Lovell., and J.J. Richardson, et al. 2014. *Communication Based Vehicle-Pedestrian Collision Warning System*. US8903640.
22. Strickland, R.D., M. Yuan., and S. Bai, et al. 2015. *Vehicle to Pedestrian Communication System and Method*. US20150035685.
23. Tengler, S., and R. Heft. 2014. *Vehicle On-Board Unit*. US8655543.
24. John, A., and Volpe National Transportation Systems Center. 2008. *Vehicle-Infrastructure Integration (VII) Initiative Benefit-Cost Analysis Version 2.3 (Draft)*. Washington, D.C.: Federal Highway Administration.

Prediction of Urban Rail Traffic Flow Based on Multiply Wavelet-ARIMA Model

Jie Zhu, Wei-xiang Xu, Hai-tao Jin and Hao Sun

Abstract Passenger flow forecast is one of basic reference for the design and operational management of urban rail transit and has become an important procedure in the construction of urban rail transit. In the paper, the ARIMA-Wavelet prediction model was established by analyzing the temporal characteristics of the daily passenger flow change and the principle of the ARIMA model and the Wavelet model, and it was used to forecast the daily traffic flow of Beijing urban rail transit line1. By using the methods of threshold processing, denoising and reconstruction of the original data, it could be better to show the features of general law, then to model and predicate the results by using ARIMA time series model. The test and analysis results of daily prediction of Beijing subway line1 traffic flow, which from November 2 to November 8 in 2015, indicted that Wavelet-ARIMA model was more accurate than ARIMA model. The Wavelet-ARIMA model received better prediction results.

Keywords Urban traffic · Wavelet analysis · Threshold processing · Data denoising · ARIMA

J. Zhu · W. Xu (✉)
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
e-mail: wxxu@bjtu.edu.cn

J. Zhu · H. Jin · H. Sun
Beijing Transportation Information Center, Beijing 100161, China

J. Zhu · H. Jin · H. Sun
Beijing Key Laboratory of Comprehensive Traffic Monitoring Service
(综合交通运行监测与服务北京市重点实验室), Beijing 100161, China

Introduction

In recent years, with the daily increase of the requirement of citizens' trip and the sharp increase in the number of urban vehicles, resulting in increasing pressure on urban traffic. Urban rail transit plays a quite backbone role in the form of transportation as one of urban transport systems and has the advantage of huge capacity, low pollution, high speed, arrival punctuality, security, and high efficiency, and it also has the influential role in alleviating traffic congestion, adding to the economic vitality of districts along rail transit, promoting the urban sustainable development, and improving the vitality and image of city. The urban rail transit in china arose late, but its tendency is very powerful in recent years. It has been in Beijing, Shanghai, and other 40 cities completed or under construction of urban rail transit lines in operation. However with the improvement of the urban rail transit network and increasing traffic flow, rail transit is facing serious congestion problems. This gives the work of operation, organization, and management of relevant departments raised new issues. And the accurate prediction of traffic flow to enter and exit station is the key to solving the problem [1].

Prediction of urban rail traffic flow is an important foundation for project approval and feasibility research report. And it is also a key basis for the entire plan and design of rail traffic system. Traditional prediction of traffic flow is mainly in consideration of the social and economic development strength of future urban and the strength of population, and use characteristics of such factors as the basis to predict the total traffic flow, which can be produced and attracted by the scope of the various stations. The commonly used method includes regression analysis, growth rate method, and cross classification analysis method [2]. Yao considering the impact of personal property, travel features, and the own characteristics of rail traffic, established the attractive model of trace site [3]. Zhirui Guang used the method of fuzzy clustering tried to separate from the specific economic data and land use numerical indicators to predict the amount of import and export station [4]. But the accuracy of the method is affected by the subjective judgment of the sample size and the type of the site.

In order to overcome the defects in existing research, we use the time series model for statistical prediction. Time series analysis is a statistical method for dynamic data processing [5]. This method is based on stochastic process theory and mathematical statistics method to study the statistical regularity which complied by random data sequence. It has the characteristics of simple structure, high-modeling speed and high accuracy of prediction. Currently, time series analysis is widely used in commercial, meteorology, agriculture, biological sciences, and other fields. Its essential characteristic is the mutual dependence between adjacent observation data; therefore, it has great practical application value. Wavelet analysis [6] method

is developed on the basis of the analysis of Fourier; it is a milestone in the development history of Fourier. And it is also a new mathematical tool for analyzing time series. Through denoising the original data of time series by wavelet denoising theory, it makes noise data to better reflect the real changes, then reestablishing time series model, and it can achieve better prediction effect and practicability.

Wavelet-ARIMA Model Building Process

For the establishment of the time series ARIMA $(p, d, q) \times (P, D, Q)$ product model, the sequence can be stabilized by differential processing. And the current method of data processing can be divided into two methods: time domain and frequency domain; in this article, the time-domain method is adopted. And SPSS software is used for drawing and computing.

Step 1: Using MATLAB tools for data wavelet decomposition, reconstruction, and denoising.

Step 2: Drawing the sequence diagram based on the denoising data. And observing whether it has the law of seasonal variation and estimating the approximate period of S, then to determine the cycle of S by the sequence of the autocorrelation function (ACF) and the test of partial autocorrelation function (PACF). In practical applications, the seasonal difference D of seasonal time series is generally not more than one order.

Step 3: Differencing the data with seasonal, then to observe whether the sequence diagram with differences is stable or not. If not stable, it should be non-seasonal difference, the number (D) of differences should be determined with the stability of the sequence diagram after differenced.

Step 4: **Pattern Recognition**. Determining the corresponding order of ARIMA model. For mixed models, the values of order P, q, P, and Q are usually taken as 0, 1 or 2. Therefore, the order of the $(P, q) \times (P, Q)$ can be defined as the choice of different combinations of 0, 1, or 2. However, the value of P, q, P, and Q cannot be zero at the same time, and hence the preferred value of P, q, P, and Q as the order of the model through the BIC criterion is chosen.

Step 5: **Parameter Estimation**. The estimated value of all the parameters in the model is estimated by the maximum likelihood estimation. ARIMA module of SPSS software is used to calculate and get the estimated value directly.

Step 6: **Diagnostic test**. Drawing ACF and PACF diagram for residual series of the $(P, d, q) \times (P, D, Q)$ product model. If the residuals autocorrelation function and partial autocorrelation functions are within the confidence

limits, indicating that the residuals are randomly distributed, and the established model is appropriate.

Step 7: Predicting according to the model. The modeling process is shown in Fig. 1.

Introduction of Principle

Denoising Principle Based on Wavelet Analysis

(1) Principle of wavelet threshold denoising

A one-dimensional signal model with noise can be expressed as follows:

$$s(k) = f(k) + \varepsilon.e(k) \quad k = 0, 1, \ldots n - 1 \tag{1}$$

Among them, $s(k)$ is a signal with noisy, $f(k)$ is a useful signal, and $e(k)$ is the noise signal. Here, $e(k)$ is considered to be a gauss white noise, which is usually expressed as a high-frequency signal. But $f(k)$ is usually a low-frequency signal in engineering practice, or it is some relatively stable signals. Therefore, it can be denoising which proceeds as follows. Firstly, the original signal is decomposed by wavelet decomposition method (as shown in Fig. 2), and the large scale approximation and a series of detail parts of the signal are obtained by layer by layer wavelet decomposition. In Fig. 2, CA3 is called an approaching signal or a smooth

Fig. 1 Block diagram of building process
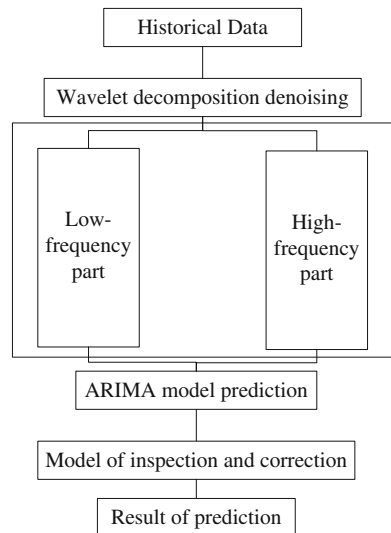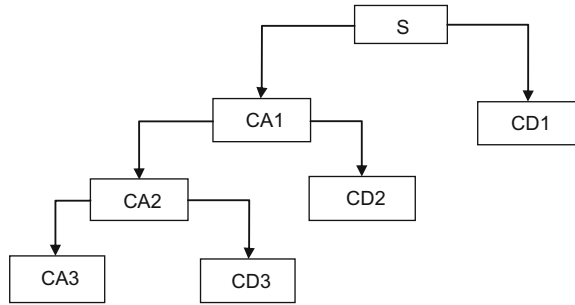
signal, and it corresponds to the low-frequency part of the signal; CD1, CD2, and CD3 are called the detail signal, which correspond to the high-frequency component of the signal. And the noise part is usually included in the CD1, CD2 and CD3; thus, the wavelet coefficients can be processed, then the signal can be reconstructed to achieve the purpose of noise elimination. The purpose of $f(k)$ denoising is to suppress the noise section in the signal. Thereby, it can restor the real signal $f(k)$ from the signal $s(k)$.

After the discrete wavelet transformation of the observed signal $s(t) = f(t) + n(t)$, from linear properties of wavelet transform, the wavelet coefficients $w_{j,k}$ obtained by decomposition are still composed of two parts. One part is the signal $s(k)$ corresponding to the wavelet coefficient $w_{s(j,k)}$, denoted as $u_{j,k}$; the other part is the signal $n(t)$ corresponding to the wavelet coefficient $w_{n(j,k)}$, denoted as $v_{j,k}$;

(2) Steps of wavelet threshold denoising

Generally, the denoising process of one-dimensional signal can be divided into the following three steps:

a. Firstly, the signal $s(k)$ with noise is transformed by disperse wavelet transform (DWT), and the wavelet coefficients of each scale are obtained.
b. Then it is made to process the wavelet coefficients of each scale by threshold processing, and obtain the estimated wavelet coefficients $\hat{w}_{j,k}$, then it is subjected to make $\left| w_{j,k} - \hat{w}_{j,k} \right|$ as small as possible.
c. Using $w_{j,k}$ to reconstruct the wavelet, the signal $s(\hat{k})$ is obtained as the estimated signal of signal $s(k)$, which is the signal after denoising.

(3) Processing of wavelet threshold denoising

Multi-resolution method can be used to characterize the non-stationary signal. Therefore, noise can be removed in different resolutions according to the distribution of the signal and the noise. After wavelet decomposition of the signal, the noise signal is generally in the high-frequency details, then it is subjected to process the wavelet coefficients of decomposition and achieve the purpose of noise reduction by wavelet reconstruction. The correlation of the signal can be removed

by wavelet transform; the noise has the trend of whitening after transforming, so the wavelet domain is more conducive to denoising than the time domain. In this paper, the observed value y_i ($i = 1, 2, 3, \ldots, n$) is transformed into wavelet domain by DWT, after that, a wavelet coefficient sequence x_i ($i = 1, 2, 3, \ldots, n$) is obtained, and then it is subjected to model and analyze after denoising.

Product Model of ARIMA

Time series analysis is a statistical analysis method, which is based on the mutual dependence between time series data and sequence, and it establishes the mathematical model by curve fitting and parameter estimation. The change of time series can be divided into four kinds, which are trend change, cycle change (seasonal variation), cycle change, and random fluctuation. Autoregressive moving average (ARMA) model is also known as Box-Jenkins (B-J) method, which is a commonly used method of time series prediction, and it is the result of the combination of autoregressive model and moving average model.

The general form of the ARMA model is as follows:

In the formula,

$$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \cdots - \phi_p y_{t-p} = \varepsilon_t + \theta \varepsilon_{t-1} + \theta \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q} \quad (2)$$

$y_t - \phi_1 y_{t-1} - \phi_2 y_{t-2} - \cdots - \phi_p y_{t-p}$	represents the autoregressive part of the model;
$\varepsilon_t + \theta \varepsilon_{t-1} + \theta \varepsilon_{t-2} + \cdots + \theta_q \varepsilon_{t-q}$	represents the moving average of the model;
$\{y_{t-p}, \ldots, y_{t-2}, y_{t-1}, y_t\}$	represents the time series data;
$\{\varepsilon_{t-q}, \ldots, \varepsilon_{t-2}, \varepsilon_{t-1}, \varepsilon_t\}$	represents the random interference sequence;
$\{\phi_1, \phi_2, \ldots, \phi_p\}$	represents the autoregression coefficient; and
$\{\theta_1, \theta_2, \ldots, \theta_q\}$	represents the moving average coefficient;

P and q are the non-negative integers, which indicate autoregressive and moving average orders. When there is a trend in time series data, the sequence can be stabilized by differential processing, it is known as quasi-stationary sequence, and the corresponding model is ARIMA (p, d, and q) model, and the model is as follows:

$$\begin{cases} \phi(B)\nabla^d y_t = \psi(B)\varepsilon_t \\ E(\varepsilon_t) = 0, \mathrm{Var}(\varepsilon_t) = \sigma_g^2, E(\varepsilon_t \varepsilon_g) = 0, s \neq t \\ Ey_s \varepsilon_t = 0, \forall s < t \end{cases} \quad (3)$$

In the formula,

d	represents the order of the difference in the process of stabilization;
B	represents delay operator;
$\nabla^d = (1 - B)^d$	represents d order difference;
$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B_p$	represents p order self regression coefficient polynomial;
$\Theta(B) = 1 - \theta_1 B - \cdots \theta_1 B_q$	represents autoregression coefficient; and
$\{\theta_1, \theta_2, \ldots, \theta_q\}$	represents q order moving average coefficient polynomial

When the time series data both have tendency and periodicity, the sequence of the time series can be stabilized by the period of difference and seasonal difference; the quasi-stationary sequence can be fitted by ARIMA $(p, d, q)(P, D, Q)_S$ model (product model of ARIMA). Wherein, D represents the order of seasonal difference, P and Q represent the autoregressive order and moving average order, and s represents seasonal cycle.

Model formula is as follows:

$$\phi(B)\phi_s(B)\nabla^d\nabla_S^D y_t = \Theta(B)\Theta_S(B)\varepsilon_t \tag{4}$$

In the expression,

$\Theta(B) = 1 - \theta_1 B - \cdots - \phi_q B_q$	represents q order moving average coefficient polynomial;
$\phi(B) = 1 - \phi_1 B - \cdots - \phi_p B_p$	represents p order self regression coefficient polynomial;
$\Theta_S(B) = 1 - \theta_1 B^S - \cdots - \phi_Q B^{QS}$	represents QS order moving average coefficient polynomial; and
$\phi_S(B) = 1 - \phi_1 B^S - \cdots - \phi_p B^{PS}$	represents PS order self regression coefficient polynomial.

Experimental Analysis

This paper takes the passenger flow of Beijing subway line1 as an example (January 1, 2010 to November 1, 2015) to build the model.

Missing Data

Data sequence analysis generally requires the data sample period is intact, but in reality, various reasons may lead to missing values. Data missing values are

Fig. 3 Data processing figure
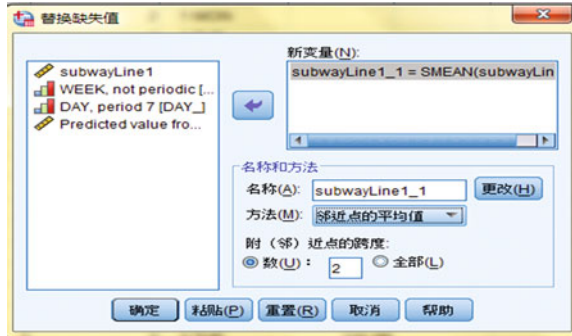
Fig. 4 Original sequence figure
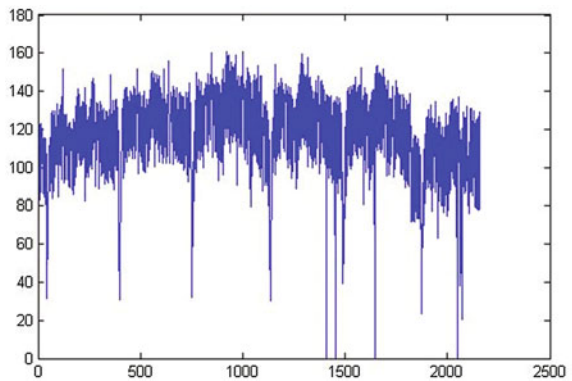
generally replaced with appropriate methods, but not simply deleted. In this paper, we use the software of SPSS to deal with missing values. This paper uses the method of adjacent point average.

This paper has selected span of proximal point as k, and selected k data points, respectively, before and after the missing date. The average value of these $2k$ data points is the missing data value. As shown in Fig. 3, the k is 2.

Wavelet Analysis

(1) According to no missing data, make the original time sequence diagram, as shown in Fig. 4.

(2) Decomposition and reconstruction

The original data with DB4 in MATLAB to do the four-layer decomposition. After decomposition, as shown in Fig. 5.

(3) Denoising analysis

The threshold means can produce a corrective action to the minimum input value in the automatic control system or cause the lowest stress reaction in tissue stimulation. Soft threshold refers to the range of detection data prescribed by the
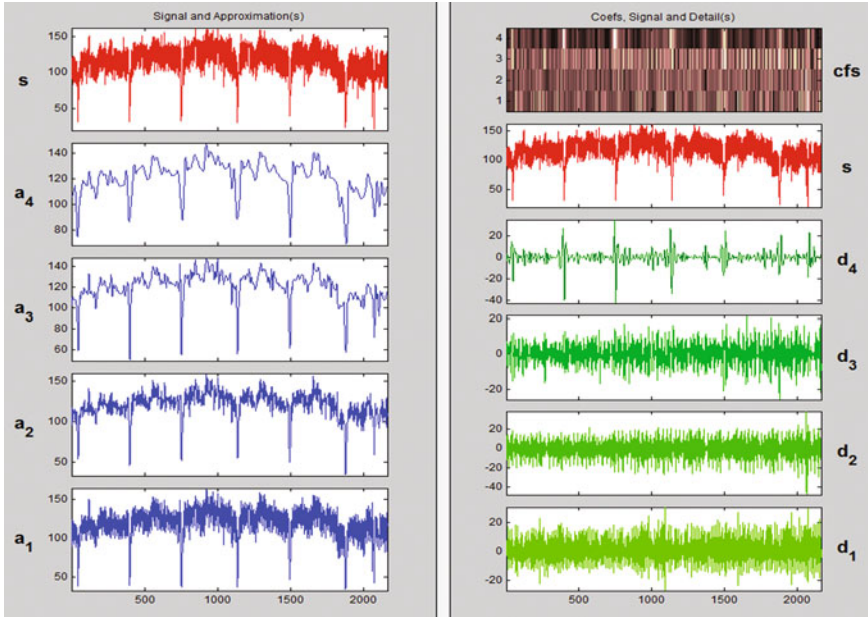
Fig. 5 Decomposition figure

threshold. This paper uses the rigorous SURE threshold method, but the data uses soft threshold processing, and the structure uses unscaled white of noise.

After several tests, the threshold settings are as follows: layer 1 threshold 16.557, layer 2 threshold 31.862, layer 3 threshold 19.931, and layer 4 threshold 45. As shown in Fig. 6.

Original and denoised signals figure as shown in Fig. 7.

ARIMA Model

(1) Cover periodicity

Figure 8 shows the time sequence diagram of traffic of subway line1 in Beijing after denoising.

According to the graph, we can see that the sequence has obvious periodicity and the period is 7 days.

(2) Determine difference

In general, the seasonal difference of the periodic sequence is not more than one order [8]. In Figure 7, (1), (2), and (3) are the diagram of $D = 1$ and $d = 0$, $d = 1$, and $d = 2$. In Figure 9, (3) is more stable, so the non-seasonal difference $d = 2$.

Take p, q, P and Q values as 0, 1, or 2 with different combinations, of course, not both zero, then by the BIC criterion, BIC value model is the smallest, and choose
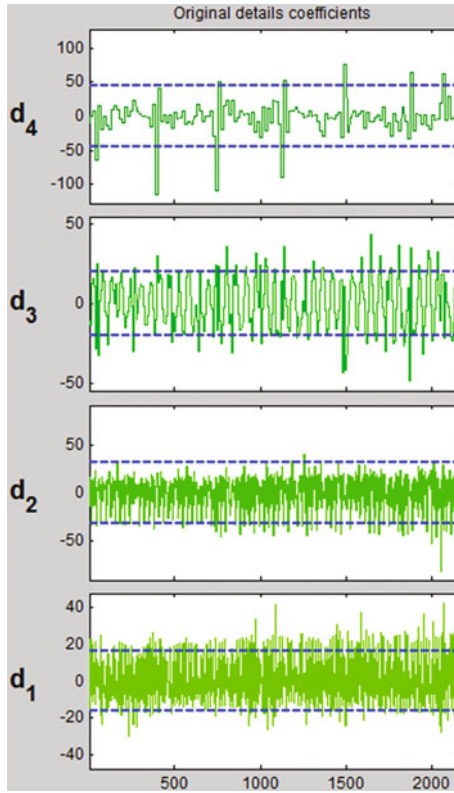
Fig. 6 Denoising process figure
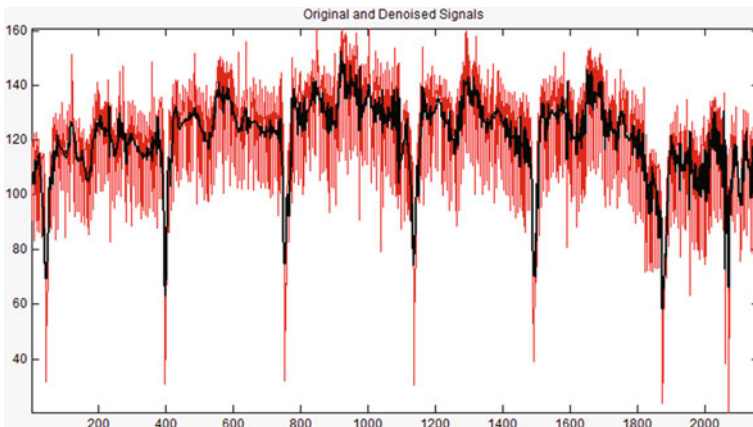
Fig. 7 Original and denoised signals figure
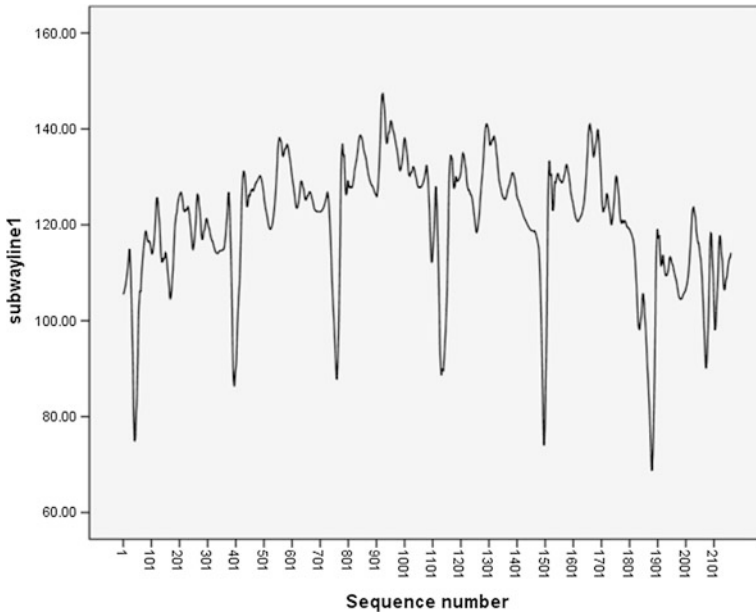
Fig. 8 Denoised signals figure

good *p, Q, P,* and *Q* values as the order of the model, by comparison, finally get a better product ARIMA $(1, 1, 1) \times (0, 1, 1)_7$ model.

(3) Parameters value

The parameter values are estimated by the SPSS ARIMA module (as shown in Table 1). The parameter values are taken into the formula (4) and get the formulas (5) and (6).

$$(1 - \phi_1 B)(1 - B_7)y_t = (1 - \theta_1 B)(1 - \theta_7 B_7)\varepsilon_t \tag{5}$$

$$\begin{aligned} y_t = & y_{t-7} + 0.452y_{t-1} - 0.452y_{t-8} + \varepsilon_t - 0.912\varepsilon_{t-1} \\ & - 0.831\varepsilon_{t-7} + 0.912 \times 0.831\varepsilon_{t-8} \end{aligned} \tag{6}$$

(4) Model checking

Figure 10 is the graph of model fitting. From the graph, it can be seen that the ARIMA model has a better fitting effect.

The R^2, MAPE, and normalized BIC were 0.871, 3.416%, and 13.121, respectively. It is also shown that the model has a good fitting effect and higher prediction accuracy.

And, autocorrelation and partial autocorrelation function for residual sequence were shown in Fig. 11. Figure 11 shows ACF and PACF were over two times the

Fig. 9 After the difference
sequence figure ($D = 1$ $d = 0$,
$D = 1$ $d = 1$, $D = 1$ and
$d = 2$)
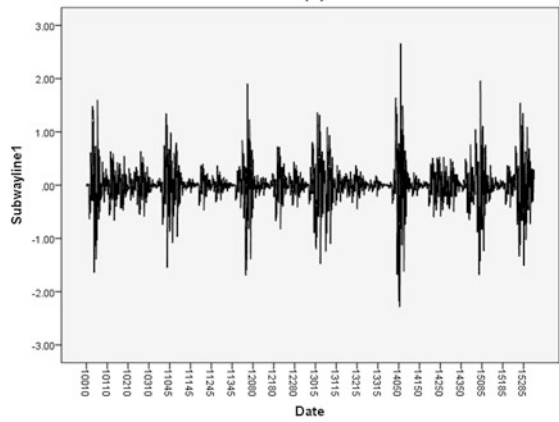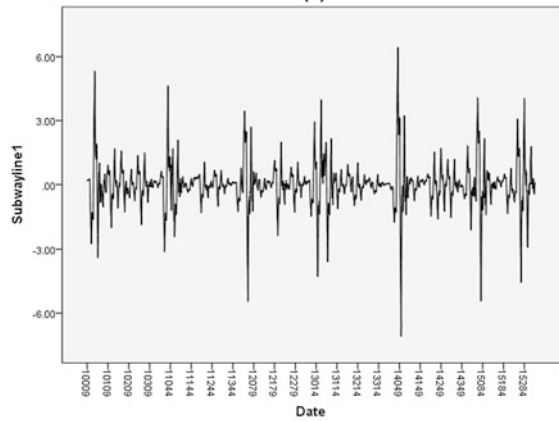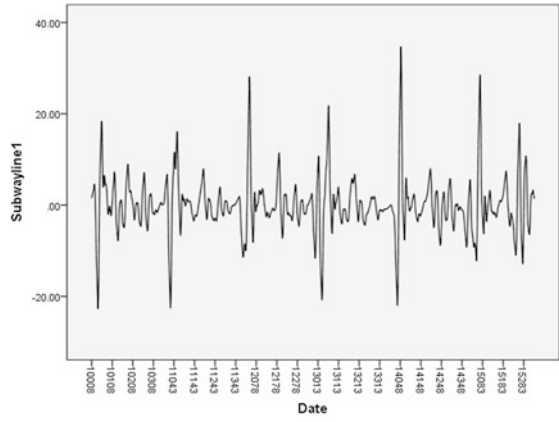
Table 1 Model parameter table

					Estimated	SE	t	Sig.
Subway line1 model 1	Subway line1	No transformation	AR	Lag 1	0.452	0.036	12.413	0.000
			MA	Lag 1	0.912	0.066	20.657	0.000
			MA, seasonal	Lag 1	0.831	0.099	48.372	0.000
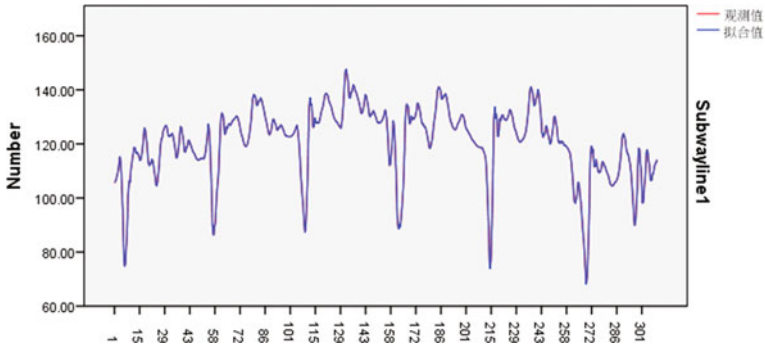
Fig. 10 Fitting figure
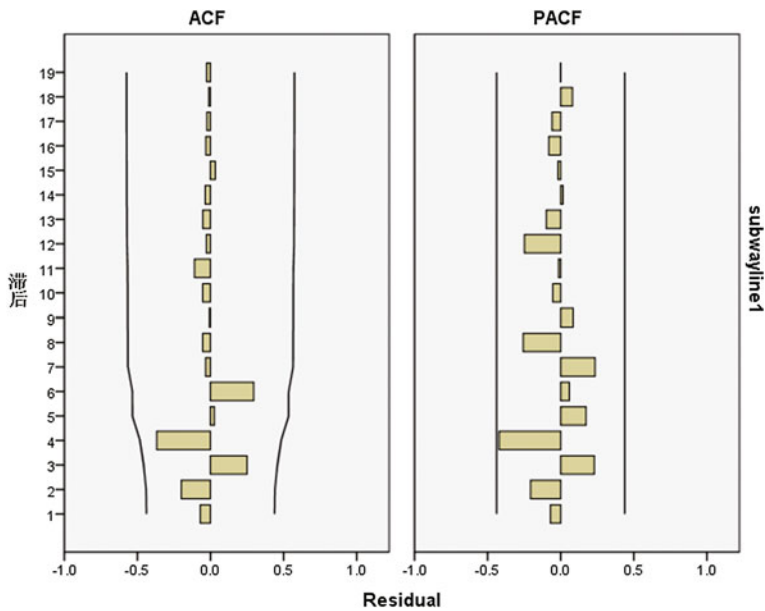
Fig. 11 Diagrams of autocorrelation and partial autocorrelation function for residual sequence

standard deviation. It is a stationary series. Therefore, the model has passed the parameter significance test and residual white noise test.

Model Validation

This paper used the ARIMA model and ARIMA-Wavelet model to predict the passenger flow of Beijing subway line1 from November 2 to November 8 in 2015, and used mean absolute percent error (MAPE) as the evaluation index of prediction results [8]. Its formula is given below:

$$\text{MAPE} = \left| \frac{(y_i - \hat{y})}{y_i} \right| \times 100, \quad (i = 1, 2, 3...n) \tag{7}$$

In this formula, y_i is the true value, and the $\hat{y}$ is the predicted value. The forecast of the two models is shown in Table 2. The MAPE is shown in Fig. 12.

The MAPE is shown in Fig. 12.

According to Fig. 12, two models with Beijing subway line1 which from November 2 to November 8 in 2015 had different forecast results. Wavelet-ARIMA prediction results of the prediction error are smaller than ARIMA model (Wavelet-ARIMA was 4%, ARIMA was 8%), so, Wavelet-ARIMA model has a better effect.

Table 2 Prediction results for the two models

Date	ARIMA-Wavelet predictive value	ARIMA predictive value	Real value	Wavelet-ARIMA error rate (%)	ARIMA error rate (%)
November 2, 2015	98.8	100.8	96	2.90	5
November 3, 2015	127.6	130.2	121.5	5	7
November 4, 2015	122.5	108.5	119.89	2.20	9.50
November 5, 2015	109.2	123.2	117.08	6.70	5.20
November 6, 2015	126.5	110.5	123.03	2.80	10
November 7, 2015	87.3	91.6	82.11	6.30	11.50
November 8, 2015	84.7	86.7	79.51	6.50	9
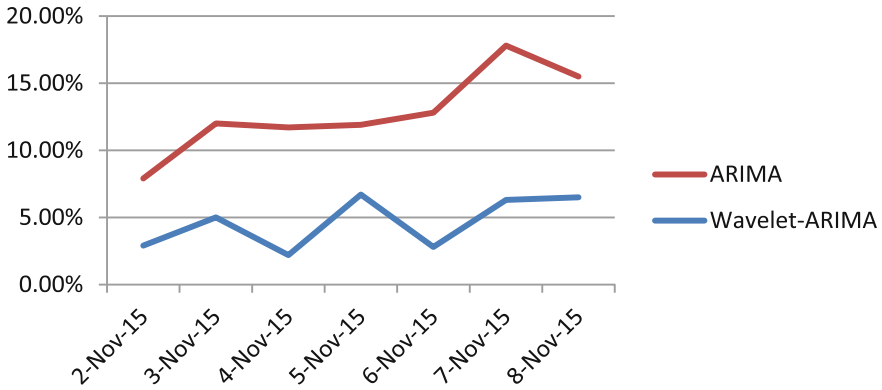
Fig. 12 Mean relative absolute percentage errors of the prediction results for the two models

Conclusions

This paper is based on Wavelet analysis model, ARIMA model, and traffic flow data (January 1, 2010 to November 1, 2015) to establish an ARIMA-Wavelet prediction model. By using the methods of threshold processing, denoising, and reconstruction of the original data, it makes the data better to show the features of general law, then predicts the results by using the ARIMA model. The test and analysis results of daily prediction of Beijing subway line1 traffic flow, which from November 2 to November 8 in 2015, indicted that Wavelet-ARIMA model is more accurate than ARIMA model. The prediction inaccuracy of Wavelet-ARIMA was 4% and the ARIMA model was 8%. The Wavelet-ARIMA model receives better prediction results. Next, we will consider the effects with weather, working day, non-working days, and other factors on the time series, and continue to study the model. The purpose is to improve the prediction level.

References

1. Cai, Changjun, Enjian Yao, Meiying Wang, and Yongsheng Zhang. 2014. Prediction of urban railway station's entrance and exit passenger flow based on multiply ARIMA model. *Journal of Beijing Jiaotong University* 38 (2): 135–140.
2. Zhang, Ning, Xia-fei Ye, and Jian-feng Liu. 2015. The impact of land use on demand of urban rail transit. *Urban Transport of China* 8 (3): 23–27.
3. Yao, Liya. 2012. *Study on the rail transportation demand forecast method based on disaggregate model*. Beijing: Beijing University of Technology.

4. Guang, Zhirui, Enjian Yao, and Yongsheng Zhang. 2012. Prediction of urban railway station's entrance and exit passenger flow based on fuzzy clustering analysis. *International Conference on railway Engineering*: 70–75.
5. Zhu, Bangzhu, and Yiming Wei. 2015. Carbon price for ecasting with a novel hybrid ARIMA and least squares support vector machines methodology. *Omega* 41: 517–524.
6. Salhi, Lotfi, Mourad Talbi, Sabeur Abid, and Adnane Cherif. 2011. Performance of wavelet analysis and neural networks for pathological voices identification. *International Journal of Electronics* 98 (9): 1129–1140.
7. Moshkbar-bakhshayesh, Khalil, and Mhammad B. Ghofrani. 2014. Development of a robust identifier for NPPS. Transients combining ARIMA model and EBP algorithm. *IEEE Transactions on Nuclear Science* 8 (61): 2383–2391.
8. Zhao, S.Z. 2010. A new approach to the prediction of passenger flow in a transit system. *Computers and Mathematics with Application* 61 (8): 1968–1974.

Simulation of Rural Vehicle Emissions Using Instantaneous Emission Model

Lijun Hao, Peng Yue, Xin Nie, Jianwei Tan and Yunshan Ge

Abstract Rural vehicles (RVs) are widely used in China rural areas and are the major sources of NOx and black carbon pollution. In order to obtain the accurate emission factors and emission inventories for RVs, the real driving emissions of several tri-wheel rural vehicles (3-W RV) were measured using portable emission measurement system (PEMS), and an instantaneous emission model was built based on the tri-wheel rural vehicle emission data and used to predict the hot stabilized 3-W RV emissions. For the selected driving trip the estimated emissions agree well with the experimental results. The instantaneous emission model can be applied to estimate 3-W RV emission characteristics for any driving pattern or driving cycle specified by the model user, and therefore the 3-W RV emission factors can be evaluated without the need of real driving emission testing.

Keywords Tri-wheel rural vehicle · PEMS · Emission simulation · Emission factor

Introduction

The rural vehicles (RVs), which are widely used in china suburban and rural areas, are usually classified into three wheelers (3-W RVs) and four wheelers (4-W RVs), and are mainly powered by diesel engine for the sake of low cost and high reliability. RVs are not allowed to drive in some urban areas and highways due to their slow speed and high emission levels. Like other diesel vehicles, the most serious pollutants from RVs are NOx and diesel particles. Especially the diesel particulate matters contain a large portion of black carbon and contribute significantly to the

L. Hao (✉) · J. Tan · Y. Ge
School of Mechanical Engineering,
Beijing Institute of Technology, Beijing 100081, China
e-mail: haolijun@bit.edu.cn

P. Yue · X. Nie
China Beijing Environment Exchange (CBEEX), Beijing 100011, China

absorption of solar radiation as an important anthropogenic aerosol component [1]. Sperling et al. [2] investigated the situation of RVs in China and estimated the total amount of emissions from RVs in 2002, and concluded that the Chinese RVs are high emitters of smoke and NOx compared to the regular vehicles. Yao et al. [3] measured the real road CO, HC, NOx, and PM emissions of RVs using a portable emission measurement system (PEMS) in China, calculated the emission factors of RVs using the International Vehicle Emission (IVE) model [4], and compared with those of light-duty diesel trucks (LDDTs). Xu et al. [5] measured CO, HC, NOx, and PM emissions of 10 3-W RVs using PEMS on national roads and rural roads, and concluded that most of the emission peaks occur under acceleration condition and the average emission factors of CO, HC, NOx, and PM on rural road are higher than that on national road. Furthermore, the emission pollutants of RVs contain high proportions of aromatics and carbonyls, which have significant impact on ozone formation [6]. Therefore, the huge population of RVs has a large contribution to the total national air pollution. In order to develop and evaluate the RV emission control legislation and policy in China, accurate emission factors and emission inventories for RVs need to be obtained, and further research work is needed to investigate the RV emission inventories at all levels.

Vehicle emissions are dependent on many parameters including engine type, vehicle weight, exhaust after-treatment technology, fuel type, driving pattern, road gradient, environmental conditions, and the level of vehicle maintenance. All these influencing factors make it necessary to utilize emission models to estimate the vehicle emissions and support the compilation of regional or national vehicle emission inventories. Various emission models were developed to predict vehicle emissions, such as COPERT [7], MOBILE [8], MOVES [9], CMEM [10], ARTEMIS [11], VERSIT [12], etc. These models are different in terms of computational requirements and the model input data required. Most of the vehicle emission models were developed for predicting automobile emission factors, and few works have been done for the estimation of RV emissions. Based on the 3-W RV emission data of CO, HC, NOx, and PM measured under real driving conditions, an instantaneous emission model was developed to estimate the emission characteristics of the 3-W RV, and the predicted exhaust emission data were compared with the real driving emission data and agreed well with the experimental results for the selected driving trip.

Experimental Section

The RV emissions are usually tested by installing its engine on test bench, and tested over the legislative engine test cycles, which cannot perfectly represent the RV real driving conditions and the level of the actual RV maintenance. Therefore, it is necessary to test the rural vehicle emissions under real driving conditions.

In order to investigate the emission characteristics of the 3-W RVs under real-world driving conditions, we used PEMS to measure CO, HC, NOx, and PM emissions of 3-W RVs on national roads and rural roads.

Tested Vehicles

In our former study, we tested 10 3-W RV vehicles [5], out of which we selected 3 3-W RVs that complied with china RV second-phase emission standard GB19756-2005, and used the tested emission data to build the instantaneous emission model for 3-W RVs. Table 1 describes the specifications of the test 3-W RVs, which represent the widely used 3-W RVs in china rural areas.

Testing Equipments

PEMSs are used for measuring emissions from vehicles under real driving conditions because they are comparatively simple, reliable, and inexpensive and can be installed on a wide variety of vehicles [13]. PEMSs are designed for measuring in-use vehicle emissions under a wide range of ambient environment.

A SEMTECH-DS analyzer (Sensors. Inc) was utilized for the on-road measurement of gaseous emissions from 3-W RVs. The SEMTECH-DS analyzer uses a non-dispersive infrared (NDIR) unit for measurements of CO_2 and CO, a heated flame ionization detector (HFID) for THC, and a non-dispersive ultraviolet (NDUV) module for NO and NO_2 measurements [14, 15]. A global positioning system receiver was utilized to obtain latitude, longitude, and altitude of vehicle's position, and the bus travel speed was calculated from these values. Temperature, pressure, and relative humidity were also measured by a weather probe and recorded for later use.

An electrical low pressure impactor (ELPI) (Dekati Ltd., Finland) was utilized to monitor the real-time number concentration (PN) and size distribution of emission particles, which will be separated into 12 channels according to their aerodynamic diameters, and each channel is connected to an electrometer current amplifier. The cut particle diameters (D50) for the 12 impactor stages are 7, 28, 55, 94, 157, 262, 384, 616, 952, 1.6, 2.4, 4.01, and 9.95 mm. The amount of particles collected in

Table 1 Specifications of 3-W RVs

Vehicle number	Manufacturer	Equipment type	Model year	Rated power/kW	Curb weight/kg
1	Foton Lovol	7YPJ	2010.12	17	1190
2	Wuzheng	7YPZ-1450A5	2010.12	16.2	1710
3	Wuzheng	7YPZ-1450PD2	2012.5	17	1770

each stage is proportional to the current value of the corresponding channel [16]. Then the measured current values are converted to an aerodynamic size distribution of the exhaust particles, and the particle mass distribution can also be obtained based on the particle density determined by the particle diameter [17].

Before each test, a series of preparing procedures were performed, including warming-up the engine and SEMTECH-DS, and calibration and leak check for the measurement devices. On-road real-time emission data for HC, CO, NOx, CO_2, and PN were measured and recorded on a second-by-second basis.

Instantaneous Emission Model

An instantaneous emission model was developed based on the second-by-second emission data (CO, HC, NOx, PM, and CO_2) of RVs as functions of vehicle speed and acceleration [18]. The structure of the instantaneous emission model (warm part only) is depicted in Fig. 1. The values of instantaneous speed and acceleration are used as inputs to the core model which consists of the base emission maps of CO, HC, NOx, PM, and CO_2 and the basic arithmetical methods. The emission maps are in the form of look-up tables with emission rates as functions of RV speed and acceleration.

The instantaneous emission model for 3-W RVs is intended to predict hot running emissions from 3-W RVs at road level and able to perform emission estimations at different geographical scales and for various traffic situations. So the RV emissions can be estimated for any driving cycle specified by the model user, and therefore new emission factors can be calculated without the need of real road emission testing.

For the emission simulation, the instantaneous acceleration is derived directly from the rate of speed change of the instantaneous speed profile, and RV speed and acceleration are used to interpolate the pollutant emission rates maps to calculate a transient pollutant emission rate. Second-by-second emission rate results are accumulated to predict the total emissions during the whole driving cycle:

$$m_{\text{tot}} = \sum m_{\text{et}} \cdot \Delta_t, \tag{1}$$

where m_{tot} is the accumulated pollutant emission mass in g, m_{et} is the mass emission rate of each pollutant in g/s at time t, and Δ_t is the simulation time step.
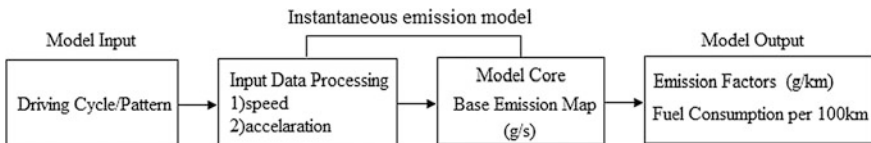
Fig. 1 Schematic of the instantaneous emission model (warm part only)

In the same way, we can calculate the vehicle total travel distance by

$$S_{tot} = \sum v_t \cdot \Delta_t,$$ (2)

where S_{tot} is the total vehicle travel distance in km, v_t is the vehicle speed in m/s at time t, and Δ_t is the simulation time step.

Therefore, we can calculate the emission factor of each emission component by

$$EF = m_{tot}/S_{tot}.$$ (3)

Results and Discussions

Figure 2 shows one of the tested hot stabilized driving trips of the Foton Lovol 3-W RV. This trip is 3.79 km long, lasts 1172s, and has an average speed of 11.67 km/h with the maximum speed of 27.9 km/h.

We used the speed-time series of this trip as the input of the instantaneous emission model, which will calculate the instantaneous acceleration and use the instantaneous speed and acceleration to interpolate the pollutant emission rates maps and calculate the transient pollutant emission rates. Next, the calculated second-by-second emission rates are accumulated to calculate the total emission of each pollutant during the whole driving trip. The calculated emission rates and the emission factors were compared with the measured emission rates and the emission factors of this real road driving trip to validate the instantaneous emission model.

The instantaneous emission model developed in this paper can be easily validated and calibrated. Any recorded 3-W RV second-by-second real road driving profile while simultaneously measuring emissions can be applied to validate the instantaneous emission model. The predicted emission rates can then be compared directly to the measured emission rates and the parameters of the model can be calibrated accordingly.
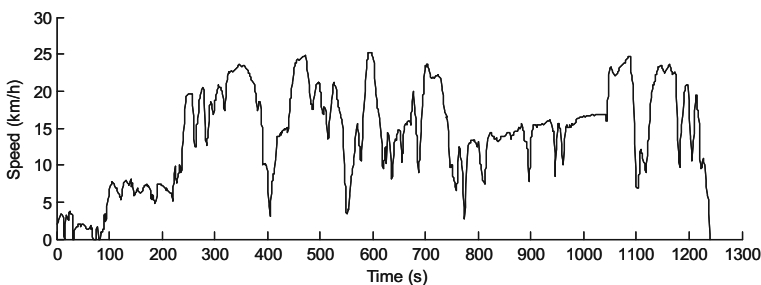
Fig. 2 Speed profile of the selected hot stabilized trip

The comparison of measured and simulated emission rates of HC, CO, NOx, and PM over the hot stabilized trip is depicted in Figs. 3, 4, 5 and 6, respectively.

As shown in Figs. 3, 4, 5 and 6, the predicted emission rates have very similar trends to the tested ones, but we can find that the predicted emission rates cannot match the measured raw data at high emission peaks and valleys, which means that our instantaneous emission model cannot predict the vehicle emissions exactly under high transient conditions. The main reason is the predicted emission rates which are derived from the emission rates maps which are composed of nodes and
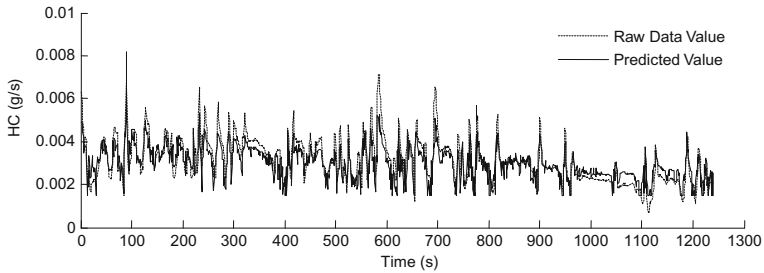
Fig. 3 Comparison of measured and simulated HC emission rates
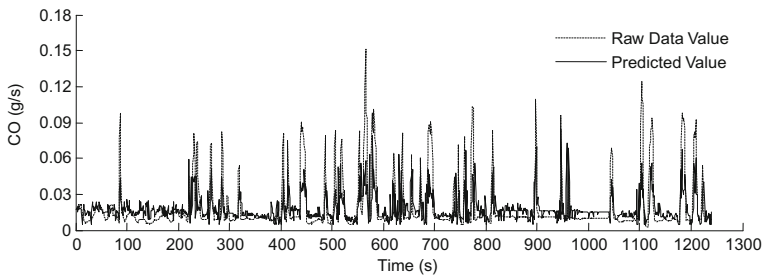
Fig. 4 Comparison of measured and simulated CO emission rates
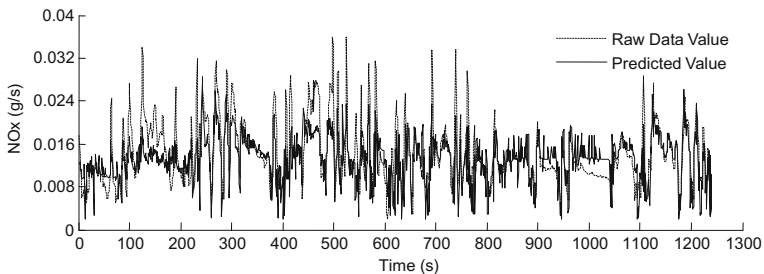
Fig. 5 Comparison of measured and simulated NOx emission rates
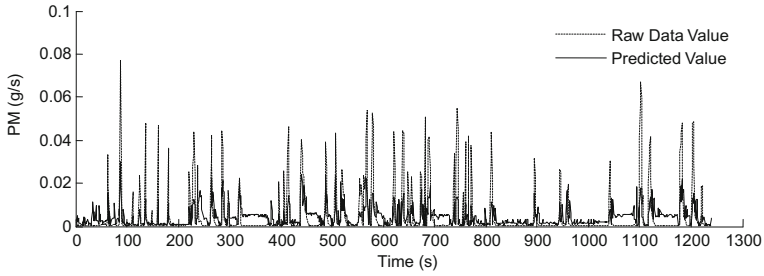
Fig. 6 Comparison of measured and simulated PM emission rates

the data of each node is the statistical average value of large amounts of raw data, so the high emission peaks have been filtered in the emission rates maps. And therefore plenty of information has to be compressed in order to get the most important topological and statistical characteristics of the tested raw data.

We also calculated the aggregated total error by comparing the calculated emission factors with the tested ones for the selected driving trip. The aggregated total error is calculated by

$$\text{Total\_error}(\%) = 100 * \text{abs}(X - Y)/Y, \qquad (4)$$

where X is the estimated data and Y is the tested data.

Table 2 shows the prediction errors of the emission factors calculated by the instantaneous emission model.

The prediction errors for the emission factors fall in the range of 3–9% for the selected driving trip. All the predicted data are smaller than the tested data. One possible reason for the prediction errors may be due to the underestimation of emission 'peaks' of the raw emission data, and there are also some other potential sources of errors, such as gear shift behavior, road gradient, and environmental conditions such as wind effect, ambient temperature, and humidity.

The agreement between the measured and calculated data for the Foton Lovol 3-W RV was acceptable. If we want to use it to predict fleet-averaged emission levels of 3-W RVs, more emission data derived from on-road test of other 3-W RVs need to be integrated into the model.

Table 2 Prediction errors for the emission factors and fuel consumption

	CO g/km	HC g/km	NOx g/km	PM g/km
Tested data	4.844	0.827	3.12	1.035
Predicted data	4.581	0.758	3.01	1.002
Total error (%)	5.43	8.34	3.53	3.19

Conclusions

In this paper, 3 3-W RV emissions were tested under real driving conditions using PEMS, and the tested emission data were used to develop an instantaneous emission model to estimate the hot stabilized 3-W RV emissions at road level. The instantaneous emission model consists of basic arithmetical methods and some emission maps built in the form of look-up tables of emission rates as functions of vehicle peed and acceleration. And the instantaneous emission model was used to estimate the emission characteristics of the Foton Lovol 3-W RV for the selected driving trip, the agreement between the predicted and measured CO, HC, NOx, and PM emissions was acceptable, the prediction errors of the emission factors were in the range of 3–9%. Using this instantaneous emission model the 3-W RV emissions can be predicted for any driving cycle specified by the model user, and therefore the 3-W RV emissions can be estimated without the need of real road emission testing.

References

1. Streets, D.G., S. Gupta, S.T. Waldhoff, M.Q. Wang, T.C. Bond, and Y. Bo. 2001. Black carbon emissions in China. *Atmospheric Environment* 35: 4281–4296.
2. Sperling, D., Z.H. Lin, and P. Hamilton. 2005. Rural vehicles in China: Appropriate policy for appropriate technology. *Transport Policy* 12: 105–119.
3. Yao, Z.L., H. Huo, Q. Zhang, D.G. Streets, and K.B. He. 2011. Gaseous and particulate emissions from rural vehicles in China. *Atmospheric Environment* 45: 3055–3061.
4. Davis N, and Lents J. Reports on the development of the IVE model and data collected in Santiago, Chile and Nairobi, Kenya. Riverside, Los Angeles: University of California.
5. Xu, Z.X., L.J. Hao, Y. Ding, H. Yin, and Y.S. Ge. 2014. Characteristics of emissions for tri-wheel vehicle under national road and rural road conditions. *Transactions of the Chinese Society of Agricultural Engineering* 30: 47–53. (in Chinese).
6. Yao, Z.L., B.B. Wu, X.B. Shen, X.Y. Cao, X. Jiang, Y. Ye, and K.B. He. 2015. On-road emission characteristics of VOCs from rural vehicles and their ozone formation potential in Beijing. *China Atmospheric Environment* 105: 91–96.
7. Ntziachristos L, and Samaras Z. 2000. Computer program to calculate emissions from road transport methodology and emission factors. Technical report No. 49. European Environment Agency, Copenhagen.
8. US EPA. 1999. A series of MOBILE6 technical reports and user's guide. Office of Mobile Sources, Ann Arbor.
9. Frey M. 1999. Methodology for developing modal emission rates for EPA's MOVES. EPA 420-R-02-027.
10. Barth M, An F, Younglove T, Scora G, Levine C, Ross M, and Wenzel T. 2001. Comprehensive modal emissions model (CMEM), version 2.02: User's Guide. University of California Riverside.

11. Andre, M., and M. Rapone. 2008. Analysis and modelling of the pollutant emissions from European cars as regards the driving characteristics and test cycles. *Atmospheric Environment* 43: 986–995.
12. Smit, R., R. Smokers, and E. Elke. 2007. A new modelling approach for road traffic emissions: VERSIT. *Transportation Research Part D* 12: 414–422.
13. Frey, H.C., A. Unal, N.M. Rouphail, and J.D. Colyar. 2003. On-road measurement of vehicle tailpipe emissions using a portable instrument. *Journal of the Air and Waste Management Association* 53: 992–1002.
14. Guo, J.D., Y.S. Ge, L.J. Hao, J.W. Tan, J.Q. Li, and X.Y. Feng. 2014. On-road measurement of regulated pollutants from diesel and CNG buses with urea selective catalytic reduction systems. *Atmospheric Environment* 99: 1–9.
15. Liu ZH, Ge YS, Johnson KC, Shah AN, Tan JW, Wang C, and Yu LX. 2011. Real-world operation conditions and on-road emissions of Beijing diesel buses measured by using portable emission measurement system and electric low-pressure. *Science of the Total Environment*. 409: 1476–1480.
16. Marjamaki, M., J. Keskinen, D.R. Chen, and D.Y.H. Pui. 2000. Performance evaluation of the electrical low-pressure impactor (ELPI). *Journal of Aerosol Science* 31: 249–261.
17. Liang B, Ge YS, Tan JW, Han XK, Gao LP, Hao LJ, Ye WT, and Dai PP. 2013. Comparison of PM emissions from a gasoline direct injected (GDI) vehicle and a port fuel injected (PFI) vehicle measured by electrical low pressure impactor (ELPI) with two fuels: Gasoline and M15 methanol gasoline. *Journal of Aerosol Science* 57: 22–31.
18. Hao, L., Chen W., Li L., Tan J., and et al. 2016. Modeling and predicting low-speed vehicle emissions as a function of driving kinematics. *Journal of Environmental Sciences*. http://dx.doi.org/10.1016/j.jes.2016.05.037.

Vehicle Routing Model and Algorithm Study for the Network of Container Transportation with Dumping Trailers Under Hard Time Window Constraint

Liupeng Jiang, Yilin Yang, Yan Zhang, Xuejun Feng and Jie Ji

Abstract The network of container transportation with dumping trailers is a typical NP-hard problem and has strict requirement on time. Therefore, vehicle routing becomes the core problem of the network. In this paper, the vehicle routing optimization model is built for the network of container transportation with dumping trailers to analyze the vehicle routing optimization problem under hard time constraint. In addition, via simulated annealing algorithm, the model is solved and finally verified via sample calculation.

Keywords Network of container transportation with dumping trailers · Vehicle routing problem · Hard time window · Simulated annealing algorithm

Introduction

Road transportation of containers is featured with short transportation distance. In addition, goods loading or unloading during container transportation is time-consuming, which will impact the vehicle transportation efficiency. Transportation with dumping trailers can reduce the stopping time of towing tractor at the loading and unloading point, thereby to improve the vehicle transportation

L. Jiang (✉) · Y. Yang · Y. Zhang · X. Feng · J. Ji
College of Harbour, Coastal and Offshore Engineering, Hohai University,
Nanjing 210098, China
e-mail: jiangliupeng@gmail.com

Y. Yang
e-mail: 2275100871@qq.com

Y. Zhang
e-mail: 2007_zy@hhu.edu.cn

X. Feng
e-mail: fxj@hhu.edu.cn

J. Ji
e-mail: 821267111@qq.com

efficiency and reduce fuel consumption and emission [1]. As a NP-hard problem, the network of container transportation with dumping trailers has core problem of vehicle routing problem (VRP) in network optimization. Under the precondition of satisfying the customers' needs, optimizing distribution route of the vehicles could realize the optimal configuration and ultimately optimize vehicle amount, total transportation mileage, and distribution time of container transportation with dumping trailers.

Study Overview

By summarizing and analyzing development of VRP, Lenstra and Rinnooy [2] and Smilowtiz et al. [3] pointed out that almost all transportation networks are NP-hard problems. Similarly, the network of container transportation with dumping trailers is also a typical NP problem. Container transportation with dumping trailers has strict requirements on time. The time requirement and VRP are both important problems for the network of container transportation with dumping trailers. Golden and Assad [4] firstly reviewed the cases involving the problem of time window and pointed out that time window constraint was an important research direction of VRP. They also proposed study framework for VRP with time window. Based on the fact whether late arrival of the vehicle to the customer will be penalized, Braysy and Gendreau [5] and Kallehauge [6] indicated that vehicle problem with time window may be divided into hard time window and soft time window problems. Kima et al. [7] studied an actual case of multi-cycle community garbage collection problem with time window and designed the simulated annealing algorithm for problem solving. Li and Lim [8] studied the goods delivery problem with time window and designed hybrid heuristic algorithm. Hashimoto et al. [9] discusses the hard time window and soft time window problems and designed tabu search algorithm aiming at soft time window problem.

With regard to solution algorithm of VRP, Solomon [10] pointed out that solution algorithm of VRP with time window was more complicated than general VRP. Solution algorithms mainly included accurate algorithm and heuristic algorithm. Examples of accurate algorithm included branching algorithm of Lysgarrd [11], column generation algorithm of Chuah and Yngling [12], and shortest route algorithm of Azi et al. [13]. Since accurate algorithm only applied to solution of small-scope problem, the bigger the problem scope was, the harder it got to obtain optimal solution with accurate algorithm. Therefore, heuristic algorithms were normally adopted. After 2000, Barysy and Gendreau [5] found that only heuristic algorithm, such as genetic algorithm, simulated annealing algorithm, and ant colony algorithm, can be used for solving large number of VRPs with time window. Caseau and Laburthe [14] proposed the tabu search algorithm to avoid tortuous

search by introducing flexible storage structure and corresponding tabu rules, thereby to realize global optimization through diversified search. Moghaddam et al. [15] studied the VRP of vehicle's independent route length and adopted hybrid simulated annealing algorithm for problem solving. Malmborg [16] analyzed the difficulty in evaluating property of the VRP solution and pointed out the advantages of genetic algorithm in algorithm implementation and global search. Adopting the genetic algorithm, they also solved the VRP with loading capacity constraint. Mazzro and Loiseau [17] adopted dynamic ant colony algorithm for solving VRP with time window and got great achievement.

Model Building

Model Overview

Vehicle dispatch plan, namely the VRP, is the core of the organization model of container transportation with dumping trailers. It is mainly optimizing vehicle routing scheme to minimize towing tractors used in transportation with dumping trailers and total travel distance of vehicles and meet the hard time requirement.

The network of container transportation with dumping trailers can be regarded as an enclosed undirected network. All vehicles depart from the dumping trailer center and finally return to the center. Firstly, before vehicle dispatching, vehicle loading is finished at each trailer loading/uploading point, the towing tractors dispatching from the freight station just arrive at the operation point to transport the loaded semitrailers to corresponding unloading point. After releasing the semitrailers, if there are containers to be transported at the unloading point, the towing tractors will proceed to operate with the trailers; if there's no container to be transported, the towing tractors will go to the next dumping trailer center to transport loaded semitrailers, and so forth until all transportation tasks are accomplished in specified time. At last, the towing tractors will return to the dumping trailer center.

Hypotheses

1. Each trailer loading/uploading point is equipped with sufficient dumping trailers, taking no account of the trailer recycling;
2. All towing tractors depart from the dumping trailer center and return to the center upon accomplishing transportation tasks;
3. Transportation with dumping tractors only takes place at the dumping trailer center, and transportation batch at the trailer loading/uploading point is inseparable;

4. All resources at the trailer loading/uploading point are sufficient, and only the cost of dumping trailers is taken into account;
5. Unloaded and fully loaded towing tractors have the same speed.

Parameter and Variable Setting

$$G = (V, E, D) \tag{1}$$

where G is the network of container transportation with dumping trailers; $V = \{v_0, v_1, v_2, \ldots v_n\}$ stands for node of the network; wherein, v_0 is the dumping trailer center and $\{v_1, v_2, \ldots, v_n\}$ are the trailer loading/uploading points; $E = \{(v_i, v_j) : v_i, v_j \in N; i \neq j\}$ stands for the arc between trailer loading/uploading points v_i and v_j; $D = d_{ij}, (i, j \in V; \ i \neq j, d_{ij} = d_{ji})$ refers to transportation distance between two trailer loading/uploading points.

$K = (k_1, k_2, \ldots, k_m)$	stands for the set of towing tractors, and M refers to the amount of towing tractors;
A	is the count of total containers transported in specified time (TEU);
L	is the count of containers each time loaded on each towing tractor (TEU);
T	is specified time;
r_a, s_a	are loading time and unloading time of task a;
q_{ij}	is the count of containers transported from point i to point j (TEU);
LT_a	is the latest time when task a is accomplished, i.e., hard time window;
d_{ij}	is the length of arc from point i to point j; $i \neq j, d_{ij} = d_{ji}$;
d_{0i}	is the distance from dumping trailer center to point v_i;
tw_i	is the trailer loading/unloading time at point v_i;
at_{ra}	is the time when towing tractor arrives at r_a;
t_{ar}	is the time when dumping trailer is loaded on towing tractor k at r_a;
t_{ae}	is the time when dumping trailer is unloaded from towing tractor k at s_a;
t_a	is the time when task a is accomplished;
P	is the penalty coefficient when towing tractor k arrives later than the time window;
Decision variable: x_{ijk}	is 0–1 variable, visiting arc of vehicle k (v_i, v_j); $x_{ijk} = 1$, otherwise, it should be 0;
y_{ak}	is 0–1 variable, the kth vehicle accomplishes the distribution task a; $y_{ak} = 1 x_{ijk} = 1$, otherwise, it should be 0.

Objective Function

Based on the model description and the hypotheses, by combining study results of Montane and Galvao [18] and Bektas and Laporte [19], the model below is obtained:

Objective function:

$$\min f(x) = (1 - \theta)M + \theta \sum_{k=1}^{M} \sum_{i=0}^{n} \sum_{j=0}^{n} d_{ij} x_{ijk} + \lambda P \tag{2}$$

In formula (2), $f(x)$ is the objective function; θ is preference parameter, indicating time preference of the network of container transportation with dumping trailers. $(1 - \theta)M$ means that the count of vehicles, one of the optimization objects, is the minimum; $\theta \sum_{k=1}^{M} \sum_{i=0}^{n} \sum_{j=0}^{n} d_{ij} x_{ijk}$ means that the total travel distance of the towing tractor, one of the optimization objects, is the shortest. λ is stage function; where $\lambda = \{0, 1\}$; if $Lt_a \geq at_{ra}$, the task a will be accomplished before the lower limit of time window; $\lambda = 0$; otherwise, $\lambda = 1$. P is the penalty coefficient, which is a large constant.

Constraint Conditions

$$\sum_{i=1}^{n} \sum_{k=1}^{m} x_{ijk} = 1, \quad \forall j = 1, 2, \ldots, n \tag{3}$$

$$\sum_{j=1}^{n} \sum_{k=1}^{m} x_{ijk} = 1, \quad \forall i = 1, 2, \ldots, n \tag{4}$$

$$A = \sum_{i=0}^{n} \sum_{j=0}^{n} q_{ij}, \quad \forall i, j \in V \tag{5}$$

$$T(M) < T \tag{6}$$

$$\sum_{k=1}^{M} y_{ak} = 1 \tag{7}$$

$$\sum_{i=1}^{n} \sum_{k=1}^{m} x_{i0k} = \sum_{j=1}^{n} \sum_{k=1}^{m} x_{0jk} = M \tag{8}$$

$$t_{ae} - t_{ar} = t_a \tag{9}$$

$$\sum_{i=1}^{n} q_i \left(\sum_{j=0}^{n} x_{ijk} \right) \leq L, \quad \forall\, k = 1, 2, \ldots, m \tag{10}$$

$$x_{ijk} = 0 \text{ or } 1, y_{ak} = 0 \text{ or } 1, \quad at_{ra} > 0, LT_a > 0 \tag{11}$$

Constraint conditions (3) and (4) ensure each customer is served by only one vehicle; constraint condition (5) means constraint on total amount of all the goods; constraint condition (6) means transportation of all the goods should be finished in specified time; constraint condition (7) means all trailer loading/uploading points are visited; constraint condition (8) means all towing tractors should dispatch from the dumping trailer center; constraint condition (9) means transportation with dumping trailers should be completed within specified time; constraint condition (10) means the total goods volume should not exceed the maximum load of towing tractor; and constraint condition (11) refers to the parameter range.

Solution Algorithm

In this paper, simulated annealing algorithm (SA) is adopted for solving VRP of the network of container transportation with dumping trailers. Inspired by similarity between solid physical annealing and combination optimization problem solving, SA is a solution algorithm for optimization problem proposed based on Metropolis principles. It is essentially a local search algorithm which avoids local optimum by introducing new mechanism, namely accepting poor solution in neighborhood with probability method to enhance diversity of the algorithm. The following are basic steps of SA:

Step (1) Firstly, obtain the initial solution x_0, initial temperature, current temperature, the maximum iterations, k, and other basic parameters.

Step (2) If the current temperature reaches the temperature when internal loop terminates, go to step (3); otherwise, please implement adjacent operation, select a neighboring solution x_j and calculate $\Delta E_{ij} = E(x_j) - E(x_i)$. If ΔE_{ij} is less than or equal to 0, or $\exp(-\Delta E_{ij}/t) > \mathrm{rank}(0,1)$, let $x_i = x_j$; repeat step (2).

Step (3) Update iterations $k = k + 1$, and current temperature according to the temperature control function, $y(t_k) = t_{k+1}$. If the algorithm termination conditions are satisfied, go to step (4); otherwise, please return to step (2).

Step (4) Export the result and terminate the algorithm. Below is the flow of applying SA to the network of container transportation with dumping trailer (Fig. 1).

Fig. 1 Algorithm flow of solving the network of container transportation with dumping trailers via SA
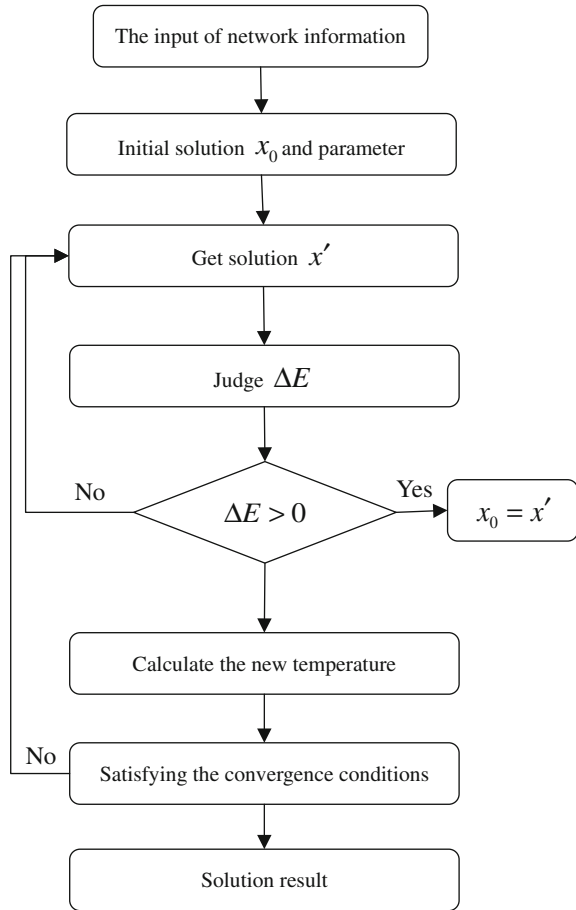
Case Analysis

In this case, Nanjing is the dumping trailer center, serving five surrounding cities (Yancheng, Nantong, Wuxi, Wuhu, and Hefei), which are regarded as the trailer loading/uploading points. Table 1 shows the distance between each two points. Being assumed as homogeneous, speed of the towing tractor is set as 60 km/h.

For a dispatched task, towing tractor should go to each point for loading/unloading dumping trailers. Time window for the task at each point is shown in Table 2. It is assumed that the towing tractor has the same unloading time at each point, i.e., 1 h.

Simulated annealing algorithm based on heuristic rule is designed according to the data above. To solve the model, relevant coefficients are shown in Table 3.

See Table 4 for assumed initial solutions.

Table 1 Distance between trailer loading/uploading points

	Nanjing	Yancheng	Nantong	Wuxi	Wuhu	Hefei
Nanjing	0	270	270	180	110	170
Yancheng	270	0	190	220	370	420
Nantong	270	190	0	140	340	400
Wuxi	180	220	140	0	240	340
Wuhu	110	370	340	240	0	150
Hefei	170	420	400	340	150	0

Table 2 Time window for each point

Point	Latest time of arrival (h)
Yancheng	9
Nantong	7
Wuxi	5
Wuhu	4
Hefei	12

Table 3 Related coefficients in model and simulated annealing algorithm

Related coefficient	Value
Initial temperature T_0	1000
Final temperature T_{end}	0.001
Internal circulation times L	1000
Temperature falling rate α	0.99
Preference coefficient θ	0.1
Penalty coefficient P	1000
Specified time T	30 h

Table 4 Initial task sequence of the towing tractor

Towing tractor No.	Task sequence
1	Yancheng
2	Nantong
3	Wuxi
4	Wuhu
5	Hefei

According to the simulated annealing principle, the optimal solution is obtained: two towing tractors are needed and the time needed for the task is 13 h. Figure 2 shows convergence of the objective function. According to Fig. 2, the algorithm converges to the extreme point 125.5 at about the 500th time; hence, the algorithm is stable.

Table 5 shows the optimal towing tractor route obtained via stimulated annealing algorithm.
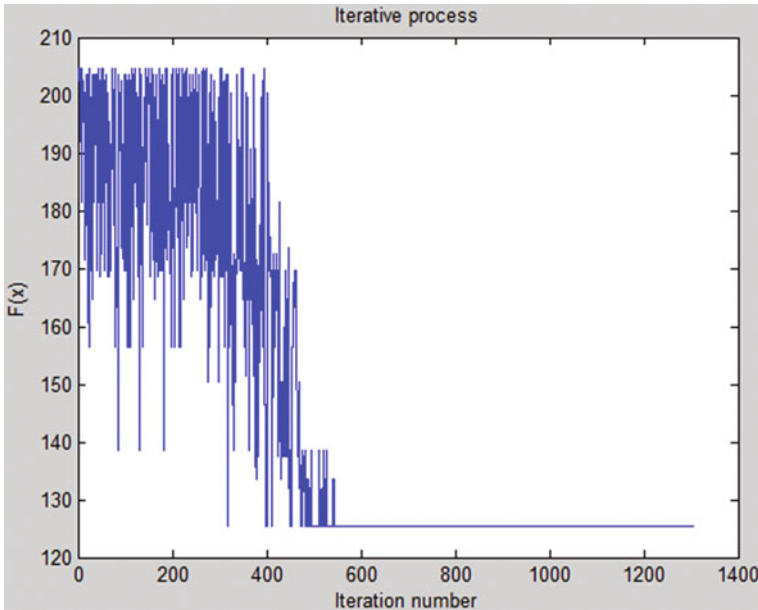
Fig. 2 Convergence graph of stimulated annealing algorithm

Table 5 Task sequence of towing tractor

Towing tractor No.	Task sequence
1	Wuhu—Heifei
2	Wuxi—Nantong—Yancheng

Conclusion

According to model and solution results above, the following conclusions can be reached: Firstly, the problem of towing tractor dispatching in network of container transportation with dumping trailers can be analyzed by building model for and solving the vehicle routing problem with time window, and a towing tractor dispatching scheme may be obtained. Secondly, simulated annealing algorithm can be used for solving vehicle routing problem with time window. It shows high computational efficiency and can get globally optimal solution. In future study, looser hypotheses will be set. Meanwhile, the problem of container transportation and collection will be the area for further study.

References

1. Wang, Z.Z., and X.Y. Liu. 2013. Comprehensive cost-benefit analysis based on the performance of drop and pull transport. In *3rd International Conference on Civil Engineering and Transportation (ICCET 2013)*, Kunming.
2. Lenstra, J.K., and K. Rinnooy. 1981. Complexity of vehicle routing and scheduling problem. *Networks* 11: 221–227.
3. Smilowtiz, K., A. Atamturk, and C. Daganzo. 2003. Deferred item and vehicle routing with in integrated networks. *Transportation Research Part E* 39 (3): 305–323.
4. Golden, B.L., and A.A. Assad. 1986. Perspectives on vehicle routing exciting new developments. *Operation Research* 34 (5): 803–809.
5. Braysy, O., and M. Gendreau. 2005. Vehicle routing problem with time windows, part 2: Metaheuristics. *Transportation Science* 39 (1): 119–139.
6. Kallehauge, B. 2008. Formulations and exact algorithms for the vehicle routing problem with time windows. *Computers & Operations Research* 35 (7): 2307–2330.
7. Kima, Byung-In, Seongbae Kimb, and Surya Sahoo. 2006. Waste collection vehicle routing problem with time windows. *Computers & Operations Research* 33: 3624–3642.
8. Li, H., and A. Lim. 2001. A metaheuristic for the pickup and delivery problem with time windows. In *13th IEEE International Conference on Tools with Artificial Intelligence*, 333–340. Los Alamitos, CA: IEEE Computer Society.
9. Hashimoto, H., T. Ybaralti, S. Imahori, et al. 2006. The vehicle routing problem with flexible time windows and traveling times. *Discrete Applied Mathematics* 154: 2271–2290.
10. Solomon, M.M. 1986. On the worst-case performance of some heuristics for the vehicle routing and scheduling problem with time window constraints. *Networks* 16: 161–174.
11. Lysgaard, J. 2006. Reach ability cuts for the vehicle routing problem with time windows. *European Journal of Operational Research* 175: 210–223.
12. Chuah, K.H., and J.C. Yngling. 2005. Routing for a just-in-time supply pickup and delivery system. *Transportation Science* 39 (3): 328–339.
13. Azi, N., M. Gendreau, and J.Y. Potvin. 2007. An exact algorithm for a single-vehicle routing problem with time windows and multiple routes. *European Journal of Operational Research* 178: 755–766.
14. Caseau, Y., and F. Laburthe. 1999. Heuristics for large constrained vehicle routing problems. *Heuristics* 5: 281–303.
15. Moghaddam, R.T., N. Safaei, and Y. Gholipour. 2006. A hybrid simulated annealing for capacitated vehicle routing problems with the dependent route length. *Applied Mathematics and Computation* 176: 445–454.
16. Malmborg, C.J. 1996. A genetic algorithm for service level based vehicle scheduling. *European Journal of Operational Research* 93 (1): 23–34.
17. Mazzro, S., and I. Loiseau. 2004. An ant colony algorithm for the capacitated vehicle routing. *Electronic Notes in Discrete Mathematics* 18: 81–86.
18. Montane, F.A.T., and R.D. Galvao. 2006. A tabu search algorithm for the vehicle routing problem with simultaneous pick-up and delivery service. *Computers & Operation Research* 33 (3): 595–629.
19. Bektas, T., and G. Laporte. 2011. The pollution-routing problem. *Transportation Research Part B: Methodological* 45 (8): 1232–1250.

A Dynamic Model of Post-disaster Search and Rescue Considering Information Uncertainty

Qianqian Liu and Qun Wang

Abstract Information moves society, so does in post-disaster SAR. Effective information management is a vital component of disaster response and relief. In the process of SAR, information is massive and constantly updated. Therefore, the decision-makers and SAR teams need to adjust the strategy, patterns, and path selection according to the dynamic information. In this paper, a grid-based distribution model considering the information as well as its uncertainty is firstly presented, in order to describe how the uncertain information influences the SAR operation, which has not been thoroughly studied by former researchers. The model incorporates with a general SAR cost, which can be used to evaluate whether the information helps to optimize the selection of SAR objectives. The model can be used to find better plans to reduce the search time and SAR costs and to help design modern simulation of SAR work to increase SAR efficiency.

Keywords Dynamic model · Search and rescue · Route choice · Information · SAR efficiency

Introduction

There is no doubt that information plays an increasingly important role in emergency management. Especially, in post-disaster search and rescue, the qualitative analysis and quantitative analysis of disasters, the location of victims and the rescue strategy selection, all of these need support of information. Some of the available information is static, such as the topographic maps, the population density, and environmental conditions of the region after disaster. But more information is

Q. Liu (✉) · Q. Wang
Business School, Hohai University, #8 West Focheng Road,
Jiangning District, Nanjing, China
e-mail: liuqq0312@foxmail.com

Fig. 1 Static and dynamic information of SAR
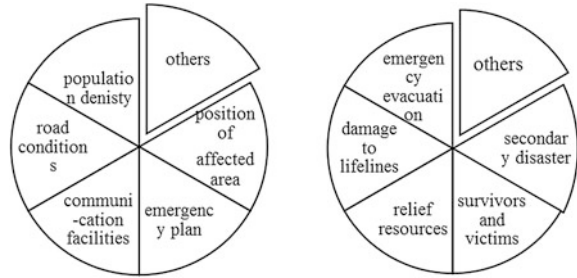
dynamic. Figure 1 is an illustration of information of SAR, in which part of information types are listed. Part of static information is shown in the left figure and some dynamic information is shown in the right figure [1–4].

Information helps to improve SAR efficiency. But only select the most useful information among vast amount information and make full use of them can achieve the goal. In SAR process, information is not only massive, but also uncertain and dynamic. In this paper, distribution models are constructed to explore the uncertainty of information in SAR, so as to better guide the distribution of rescue resources.

Post-disaster SAR is a systematic project. "Locate, search and rescue the trapped survivors" is the main aim of post-disaster SAR. The rescue effect is often influenced by rescue difficulty and relief resources. Nowadays, the researches about SAR are mainly focusing on the emergency logistics and emergency decision-making [5–12]. Due to the damaged environment, only part of the SAR resources actually comes into use. When the operational areas become larger or the conditions become more complex, the efficiency of SAR will decrease. After disasters, the decision-makers will formulate a set of relief strategies. Strategies are not invariable all the time; they will be improved with information updating and the change of disaster conditions. All instantaneous information such as the number of victims, survival rate, and SAR teams' allocation helps to improve the SAR strategy and relief work. So it is necessary to research the characteristic of information in post-disaster SAR so as to optimize SAR work.

To SAR teams, the number and location of survivors, the traffic condition and rescue time is uncertain. Information has a significant influence on SAR resources distribution. The information about the disaster environment and the distribution of survivors will raise the SAR efficiency obviously, and the survivors' waiting time will minimize. The characteristic of SAR varies in different SAR operations. In order to measure the effect of information in SAR operations, it is necessary to establish models based on SAR characteristic.

The Assumptions

The Behavior of Search and Rescue

Consider the following scene: a remote region was attacked by flood and most of the area was damaged. The area was divided into several grids and there were some survivors waiting for rescue. After disaster, only a few paths were still passable and their conditions can be observed by positioning device. Due to the damage of communication network, the rescue center could only get the approximate location of survivors through mobile phone signals. The SAR teams can learn each other's position and path selection information; however, they had to cruise to search and rescue more survivors.

Due to the terrible ruined environment and poor information management after disaster, we do not consider a central rescue management system in this paper. All teams will carry out the rescue work independently and dispersedly. The SAR work generally includes four processes as shown in Fig. 2.

After the SAR teams send the survivors back to rescue center, this round SAR work completed. Then, a new round rescue work will begin. It is worth to note that the behavior pattern of process B and D is different from process C. In process B and D, the motion trail of SAR teams is directive movement. That is to say, the teams all have clear goals before their movements. Once the goals are given, they will choose the shortest path and go to the target grid directly. Similarly, when SAR teams find survivors, they will send the survivors back to rescue center directly after basic treatment. But in process C, SAR teams do regular cruising movements.

The Basic Assumption

According to the scene described above: Suppose the rest available paths as a passable grid net $G(V, A)$, in which V is the set of nodes and A is the set of arcs (linked roads) in the network. There are a number of SAR teams searching

Fig. 2 Four processes of SAR

A. Decide the target grid g according to the instantaneous information

B. Select the shortest path and go to the target grid g

C. Cruise and search for survivors after arriving at the target grid g

D. Send the survivors back to rescue center m after basic treatment
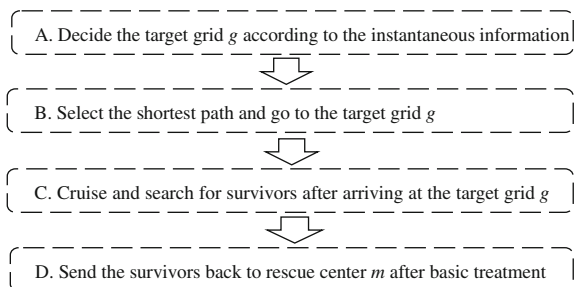
survivors in the road network, and the quantity is given for simplicity. Because of disaster, the SAR teams cannot get accurate information about the survivors in time. All survivors need to be sent to the rescue center (node m) for medical treatment. We suppose all survivors will be sent to the rescue center within the time period T, and $t \in T$.

The grid net G is divided into several grids, let g be one of them ($g \in G$). Assume that there are still some survivors in grid g at time t, and let it be $S_g(t)$. Let S be the total number of survivors in the network, then

$$S = \sum_{g \in G} S_g \tag{1}$$

Assume that the remaining passable paths are not congested, and there is no delay because of queue or danger. Let c_{gm} be the travel time from m to grid g, and c_{mg} is constant. Suppose there are N SAR teams searching survivors in the network, and all teams will go to each grid to work from the rescue center. Let $N_m(t)$ be the number of SAR teams in rescue center at time t, there is

$$N = N_m(1) \tag{2}$$

At the initial time $t = 1$, all SAR teams gather in the rescue center. Let $N_g(t)$ be the number of SAR teams searching in grid g.

As shown above, the goal of SAR is to rescue survivors as more as possible, that is

$$Z = \sum_{g \in G} S_g \cdot W_g \tag{3}$$

where W_g is the survivor's average waiting time in grid g, and Eq. (3) can be regarded as the evaluation index of SAR operations.

The Expected Search Time and Waiting Time

The Expected Search Time

Consider a SAR team that reaches grid g at time t, and its expected search time spending in finding a survivor will be derived in this section. Because the information is continuously updated during the SAR process, we assume that the SAR teams only estimate expected search time according to the information at time t. There are $S_g(t)$ survivors and $N_g(t)$ SAR teams in grid g at this moment, and only the conditions of $N_g(t) \leq S_g(t)$ is considered, which means rescue resources waste does not exist.

Suppose the survivors in the grid network are uniformly distributed. Although the exact location is not clear, we can obtain the distribution density $\rho_g(t)$ by calculation,

$$\rho_g(t) = S_g(t)/A_g \qquad (4)$$

Consider that there is a SAR team searching in the grid with speed v, so after one unit time its search distance will be $v \cdot 1$ and the probability it finds a survivor will be $v \cdot \rho_g(t)$. There are $N_g(t)$ SAR teams in the network, and the number of survivors they find after a unit time will be

$$f_{gm}(t) = N_g(t) \cdot v \cdot \rho_g(t) \qquad (5)$$

where $f_{gm}(t)$ is the number of SAR teams back to rescue center from grid g at time t. At time $(t+1)$, the SAR teams and survivors in the network reduce $f_{gm}(t)$, that is

$$N_g(t+1) = N_g(t) - f_{gm}(t) \qquad (6)$$

$$S_g(t+1) = N_g(t) - f_{gm}(t) \qquad (7)$$

Then, at time $(t+1)$, the distribution density decreases, so

$$f_{gm}(t+1) = N_g(t+1) \cdot v \cdot \rho_g(t+1) \qquad (8)$$

where $\rho_g(t+1) = S_g(t+1)/A_g$. Thus, it can be determined that at time $(t+x')$ each SAR team finds survivors, which means at time $(t+x')$, the value of $N_g(t+x')$ equals zero.

When the SAR team leaves grid g at time $(t+x)$, its search time is x. So when the SAR team reaches grid g at time t, the expected search time $ec_g^s(t)$ will be

$$ec_g^s(t) = \sum_{x=1}^{x'} (f_{gm}(t+x) \cdot x)/N_g(t) \qquad (9)$$

where α_g can be obtained by survey parameters. When $N_g(t) > S_g(t)$, part of SAR teams cannot find any survivors, so this time period will be regarded as invalid and it can be set to a certain value. Then, the expected search time can be treated similarly.

Waiting Time

The survivors' waiting time W_g can be obtained by

$$W_g = \sum_t f_{gm}(t) \cdot t / S_g \qquad (10)$$

where $\sum_t f_{gm}(t) \cdot t$ is the sum total of survivors' waiting time from rescue center m to grid g. In this way, we can get the average waiting time.

The Influence of Information on SAR

To SAR teams, the exact locations of survivors are not certain so they have to cruise independently in the target grids. As shown in Fig. 3a, there are one rescue center and three grids in an example network. The corresponding decision network is shown in Fig. 3b, which shows that at time t, each SAR team needs to choose one grid among the three grids as their target and then go to this grid through the shortest path. Consider that from the rescue center m to the target grid g, it will cost c_{mg} time. Thus, when each SAR team chooses their target grid at time t, how much time will spend to find a survivor in this grid at time $t + c_{mg}$ should be evaluated either.

Let $ec_g^s \left(t + c_{mg} \right)$ be the expected search time in grid g at time t. In other words, a SAR team is estimated to spend $ec_g^s \left(t + c_{mg} \right)$ time in finding a survivor. Then, the total search time $C_{mg}(t)$ will be the sum of the expected search time and the travel time (from rescue center m to target grid g).

$$C_{mg}(t) = c_{mg} + ec_g^s \left(t + c_{mg} \right) \qquad (11)$$

Due to the damage of disasters, for SAR teams, all the location of survivors, the traffic conditions, and rescue time are not certain. So perception error exists in search time and rescue time for SAR teams. But with the updating of information, the perception error will decrease.

Fig. 3 Illustrator network
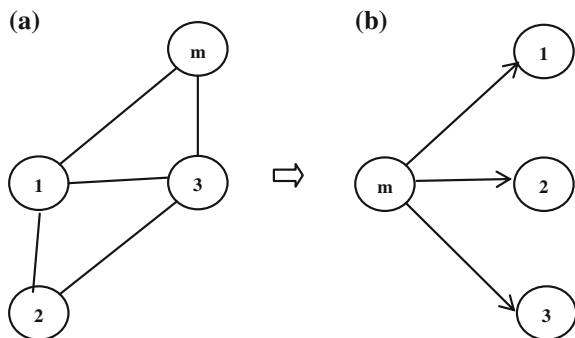
The perception error is assumed to be identically distributed with a *Gumbel* density function. According to classic models in economics and engineering, we suppose the perception error of expected search time shows Gumbel distribution for simplicity [13, 14]. Based on these behavior assumptions, the probability a SAR team goes to grid g from rescue center m and then finds a survivor eventually in this grid is specified by the following logit model

$$Pr_g(t) = \exp\big[-\gamma \cdot C_{mg}(t)\big] \bigg/ \sum_{g \in G} \exp\big[-\gamma \cdot C_{mg}(t)\big] \qquad (12)$$

where γ is a nonnegative parameter indicating the uncertainty of SAR process, and from observational data, γ can be calibrated. The value of γ reflects the degree of uncertainty on disaster environment and location of survivors. A small value of γ means less information about the survivors' location is available to SAR teams, so they have to search by cruising. On the other hand, if γ is large, more information is available and the random error is small. SAR teams will tend to minimize search time and optimize grid selection.

The Uncertain Information and Distribution Model

The Distribution of SAR Teams

All SAR teams leave for target grids from rescue center m at initial time. The total quantity is known, but their exact location is not clear. So, the SAR teams can only choose one target grid to search by cruising. $C_{mg}(t)$ is the SAR cost that one SAR team spends in grid g. The cost should be minimized so as to search as much survivors as possible. We assume each team tries to minimize the SAR cost when they choose their target grids.

Although the survivors' exact location are not clear, the SAR teams still can get some information through communication facilities and positioning device. It is assumed that the SAR teams can get some continues and instantaneous information, and they can get the distribution of SAR teams and the quantity of rescued survivors. Based on this, they can assess the search difficulty of each grid and choose their target grids.

According to the assumption above, the probability a SAR team goes to grid g from rescue center m and finds a survivor eventually in this grid is specified by the logit model as shown in Eq. (12), then to the $f_{mg}(t)$

$$f_{mg}(t) = N_m(t) \cdot \exp\big[-\gamma \cdot C_{mg}(t)\big] \bigg/ \sum_{g \in G} \exp\big[-\gamma \cdot C_{mg}(t)\big] \qquad (13)$$

In order to improve rescue success rate, each SAR team has to find survivors as much as possible in limited time, which means to minimize the rescue cost of each round. Then, the problem transforms into a classical stochastic network assignment.

The SAR conditions changes with time, and so does $c_g^s(t)$. The SAR teams need to select the target grids to search according to the changing conditions, which means, during different time periods the SAR teams from m have different target grids.

The Distribution Model

When SAR teams with survivors go back to the rescue center m, this round of rescue task finishes and a new round rescue task begins. During the whole-time period, one SAR team can perform SAR mission for many times. At time t, the total of SAR teams from m is $N_m(t)$. As the number of SAR teams at the initial time is known, so to grid g at time t $(t > 1)$

$$N_g(t) = N_g(t-1) - f_{gm}(t) + f_{mg}\left(t - c_{mg}\right) \tag{14}$$

As shown in Eq. (16), at time t, the original number of SAR teams in grid g is $N_g(t-1)$, then $f_{gm}(t)$ teams find survivors and go back to the rescue center, at the same time $f_{mg}\left(t - c_{mg}\right)$ teams reach grid g and begin to search. Note that the SAR teams spend time c_{mg} to reach grid g, so they begin to search at time $(t - c_{mg})$. The quantity of SAR teams at rescue center at time t is

$$N_m(t) = N_m(t-1) - \sum_{g \in G} f_{mg}(t) + \sum_{g \in G} f_{gm}\left(t - c_{mg}\right) \tag{15}$$

Similarly, at that moment the number of survivors is

$$S_g(t) = S_g(t-1) - f_{gm}(t-1) \tag{16}$$

A Solution Algorithm

The algorithm involves the SAR teams' selection equation can refer to the commonly used stochastic network distribution method. In fact, that this problem consists of two loops, where one loop is used to calculate the search time in a grid at time t, and the other loop is used to calculate the distribution flow. The basic idea is on the basis of initial value, iterate step by step for optimal solution. Details are as follows:

Step 1. Initialization. Let all variables $\{N_m(t), N_g(t), f_{mg}(t), f_{gm}(t)\}$ be the initial value according to the constraints and conditions. Set iteration counter $k = 0$.

Step 2. Update. Set $k = k + 1$. Compute SAR teams and survivors in each grid, $S_g(t)^{(k)}$ $N_g(t)^{(k)}$, $\forall g$ according to,

$$N_g(t)^{(k)} = N_g(t-1)^{(k-1)} - f_{gm}(t-1)^{(k-1)} + f_{mg}\left(t - c_{mg}\right)^{(k-1)}$$
$$S_g(t)^{(k)} = S_g(t-1)^{(k-1)} - f_{gm}(t-1)^{(k-1)}$$

Step 3. Compute Search time, according to,

$$c_g^s(t)^{(k)} = \ln N_g(t)^{(k)} / 2\ln(1 - \alpha_g)^{-1}$$

Step 4. Compute search cost $C_{mg}(t)^{(k)}$, $\forall g$, according to the following equation, and compute $C_m^*(t)^{(k)}$.

$$C_{mg}(t)^{(k)} = c_{mg} + c_g^s\left(t + c_{mg}\right)^{(k)} + c_g^r$$

Step 5. Assign the search teams, using classic stochastic assignment method this yields $f_{mg}(t)^{(k)}$, so that,

$$f_{mg}(t)^{(k)} = N_g(t)^{(k)} \cdot \exp\left[-\gamma \cdot C_{mg}(t)\right] \bigg/ \sum_{g \in G} \exp\left[-\gamma \cdot C_{mg}(t)\right] \qquad (17)$$

Step 6. Convergence test. If all search teams arrive at their destinations, which means $|f_{mg}^{(k)} - f_{mg}^{(k-1)}| < \epsilon$, where ϵ is a predetermined convergence tolerance, stop; otherwise, go to step 2.

Conclusion

In this paper, the SAR work is first divided into four stages, and it is pointed that the behavior mode in each stage is different. By proposing a distribution model, we described in an uncertain and constantly updated status, how the SAR teams select target zones and cruise to search survivors with dynamic information. It is found that during the SAR process, the SAR teams will adjust their targets and routes with the instantaneous information. The more information obtained, the clearer the goals will be, and the lower search time and search cost, correspondingly the higher efficiency. Of course, limitations still exist in this paper, for example the survival

rate has not been taken into account, and an illustration example is needed. As a very important variable, the survival rate and the search detail in a grid will be considered in our future research.

References

1. Masakatsu, A., Shunichi, Y., and Kenichiro, S. 2007. Disaster-response information sharing system based on cellular phone with GPS. In *Second International Conference*, OCSC 2007, 250–255.
2. Mclntosh, S.E., A. Brillhart, J. Dow, and C.K. Grissom. 2010. Search and rescue activity on Denali 1990 to 2008. *Wilderness & Environmental Medicine* 21 (2): 103–108.
3. Hung, E.K., and D.A. Townes. 2007. Search and rescue in Yosemite National Park: A 10-year review. *Wilderness & Environmental Medicine* 18 (2): 111–116.
4. Calisi, D., L. Iocchi, D. Nardi, G. Randelli, and V.A. Ziparo. 2009. Improving search and rescue using contextual information. *Advanced Robotics* 23: 1179–1198.
5. Fiedrich, F., F. Gehbauer, and U. Rickers. 2007. Optimized resource allocation for emergency response after earthquake disasters. *Safety Science* 35: 41–57.
6. Yuan, Y., and D.W. Wang. 2009. Path selection model and algorithm for emergency logistics management. *Computers & Industrial Engineering* 56: 1081–1094.
7. Nagy, G., and S. Salhi. 2007. Location-routing issues, models and methods. *European Journal of Operational Research* 177 (2): 649–672.
8. Sheu, J.B. 2007. Challenges of emergency logistics management. *Transportation Research Part E* 43: 655–659.
9. Wang, W., Liu, M., Wang, L. 2010. Dynamic optimization methods of emergency resources deployment based on Markov decision process for Wenchuan earthquake. In *2010 2nd International Workshop on Date of Conference*, Wuhan, 1–4.
10. Janis, I.L., and L. Mann. 1977. Emergency decision making: A theoretical analysis of responses to disaster warnings. *Journal of Human Stress* 3 (2): 35–45.
11. Kapucu, N., Garayev, V. 2011. Collaborative decision-making in emergency and disaster management. *International Journal of Public Administration*, 34: 366–375.
12. Yu, Z.Y., and Han, C.F. 2014. Emergency decision making: A dynamic approach. In *Proceedings of the 11th International ISCRAM Conference*. USA, 240–244.
13. David, A., W. Hensher, and H. Greene. 2003. The mixed logit model: The state of practice. *Transportation* 30: 133–176.
14. Nadarajah, S., and S. Kotz. 2006. The beta Gumbel distribution. *Mathematical Problems in Engineering* 91 (6): 689–697.
15. Yang, H., and S.C. Wong. 1998. A network model of urban taxi services. *Transportation Research Part B* 4: 235–246.

Study on Optimization and Adjustment Method of Urban Public Transport Network Based on Evolutionary Analysis

Wenyong Li, Tao Wang and Xiqi Zhang

Abstract Based on the complexity analysis of urban public transport network evolution, according to the passenger interests and construction cost restraint mechanism, using the complex network L modeling of space method, evolution model of urban public transport network is proposed. Through researching the interests of passing passenger and arriving passenger constraint mechanism, optimizing and adjusting model of urban public transit network is proposed. Simulation results show that the evolution of the urban public transport network should first consider the interest of the passengers, and then only consider the appropriate investment on construction and operation cost. In the optimization and adjustment of urban public transport routes, arrival passenger restraint mechanism is a key consideration, passing passenger restraint mechanism being the secondary factor. Compared with the traditional model in the form of variable and vector, the public transit network optimization and adjustment method are based on evolvement analysis which shows the spatial layout optimization goal of public transit network.

Keywords Urban public transport network · Evolutionary analysis · Optimization adjustment · Complex networks

W. Li · T. Wang (✉)
School of Architecture and Transportation, Guilin University
of Electronic Technology, Guilin 541004, China
e-mail: wangtao_seu@163.com

T. Wang
School of Transportation, Southeast University, Nanjing 210096, China

X. Zhang
Guilin Urban Planning and Design Institute, Guilin 541001, China

Introduction

In the face of limited space and time resources of the urban road, the traffic demand and supply becomes more and more contradictory. Promoting bus travel will ease traffic congestion, improves the utilization rate of road resources, and improves the efficiency of travel. At the technical level, the scientific urban public transport network planning, optimization, and adjustment are the key factors to improve efficiency.

However, the conventional public transport network planning is a one-way and static process; the method makes use of the data from traffic survey to forecast traffic demand, and then determines the supply to optimize public transport network. The conventional public transport network optimization study is missing from the perspective of the network and the overall structure. Actually, traffic demand and the traffic supply influence each other; there is a bidirectional and dynamic relationship. Therefore, the shortcomings of traditional public transport network planning model were reflected in this paper, and then the fundamental study of the optimization and adjustment method of urban public transport network based on evolvement analysis was carried out.

The Complexity Analysis of Evolution

Evolutionary Basis

Nodes and lines are the basic unit of the traffic network [1]; bus stops and routes are the "nodes" and "lines" of the urban public transport network. Traffic network evolution process is the process based on the nodes and lines change.

In the process of the urban public transport network evolution, bus stops and routes will change. For example, adding of a new stop or a new route, and adjusting the route [2]. The complexity of the stops and routes change determines the complexity of urban public transport network evolution study.

Complexity of Network Entitlement Evolution

Network is divided into the non-entitled network and entitled networks [1]. Non-entitled network revolution research mainly refers to the research of network topology mechanism, and on this basis to analysis the technology structure. The entitled network revolution research mainly refers to the interaction between network topology and network load factors. Relative to the non-entitled network revolution study, the entitled network revolution study can better reflect the real

network of internal and external attribute and find the real network revolution mechanism [3].

In summary, non-entitled urban public transport network can reflect the basic features of the evolution. And entitled urban public transport network to some extent, can reflect the evolution of the real property. The entitled urban public transport network is varied, such as the physical distance between the bus stations, travel time, traffic volume, and so on. Different entitlements have different meanings. The network characteristics which the entitlements reflect are also different. So in the research of the urban public transport network, the selection of the entitlements and the overall planning of the urban public transport network increased the complexity of the evolution process.

Complexity of Evolutionary Mechanism

The network evolution of single evolutionary mechanism can fully reflect the role of a single evolutionary mechanism on network evolution. The hybrid evolution mechanism of the network refers to the network with multiple evolution mechanism combination evolution, which can reflect the network evolution characteristics more realistically [4].

The evolution of urban public transport network is affected by many factors, so the evolution is the result of a variety of evolving mechanisms, belonging to the hybrid evolution mechanism. The complexity of the evolution process of urban public transport network is mainly for:

1. According to the representation of urban public transport network revolution, deducing the correlation factors of various evolution. And on the basis of this, describing the evolution mechanism.
2. According to the different development periods of the urban public transport network, requiring different levels of the network structure. Analyzing the evolution mechanism of category and determined the influence of relative degree.

Evolutionary of Models

Evolutionary of Model Construction

For the interests of passengers, the bus stations always aim to be as close to the destination as possible [5], i.e.,

$$B_p(i,j) = \frac{1/d_{ij}}{\sum_{j \in v(i)} 1/d_{ij}} \tag{1}$$

For public transportation construction and operation costs, in the premise of meeting the traffic demand, the cost of construction should be as low as possible [6, 7], i.e.,

$$B_c(i,j) = \frac{1/c_{ij}}{\sum_{j \in v(i)} 1/c_{ij}}, \quad c_{ij}(w_{ij}, d_{ij}) = \frac{0.905 d_{ij}^{1.558}}{(s_i s_j)^{0.407}}, \quad w_{ij} = K \frac{s_i s_j}{d_{ij}^2} \tag{2}$$

In Formula (1), d is the distance between two adjacent bus stations i and j. In Formula (2), s is the strength of the stop. This paper established an evolutionary model of urban public transport station network, which is based on the interests of passengers and the cost of construction and operation. The model is divided into two layers: The upper layer is the urban public transport travel demand entitled network model, and the lower is the urban public transport station network model which is constructed with the L space method [8, 9]. Urban public transport travel demand entitled network model refers to any urban public transport network in two stops of OD, which means traffic demand. The network model of urban public transport station is constrained by two evolution mechanisms: The interests of the public transport travel distance constraints mechanism and the construction and operation of the cost constraint mechanism. The evolutionary model has the following three characteristics:

1. The interests of the passengers and the impact of the construction and operating cost are took into account at the same time: In the evolution of urban public transport network, the interests of passengers are reflected in the shorter distance, but the construction and operation cost factors cannot be ignored;
2. Entitlement evolution: With the evolution of the urban public transport station network, the entitlement travel demand network will also evolve; and
3. Volume impact: With the connection of the nodes in the public transport station network, the traffic flow between the nodes will be affected.

Evolutionary Mechanism

The evolution mechanism of the model include: (1) Line growth mechanism; (2) Route traffic demand-oriented mechanism; (3) Passenger interests restraint mechanism; (4) Construction and operation cost restraint mechanism; (5) Line entitlement evolution mechanism; (6) Traffic volume impact evolution mechanism; and (7) Upper and lower network dynamic feedback mechanism. The dynamic evolution feedback mechanism of the model is shown in Fig. 1.
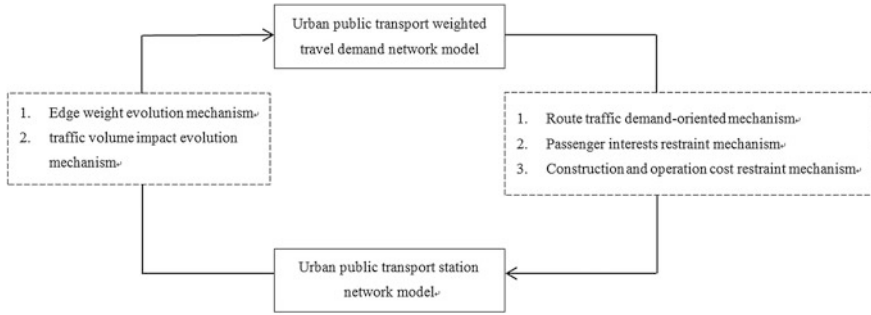
Fig. 1 Schematic diagram

Model Evolutionary Processes and Algorithms

1. Network initialization: Using the L space method to build the public transport station network;
2. Add a route: Each time step t ($1 < t < T$), in accordance with the following rules to add a route which has m stops to the urban public transport network:

   Step one: Determine the terminal station of the route. Combined with the gravity model, the first terminal is determined in the initial network, its connection probability is

$$p_{OD} = \frac{s_O s_D}{\sum_{a < b} s_a s_b} \tag{3}$$

   Step two: Build an alternative set of next stop. Using the judging principle of line trend put forward by Sui, et al. [10]:

$$\beta(S_{ti}, K_j) = \begin{cases} 1, & 0 \le \cos(S_{ti}K_j, S_{ti}K_j) \le 1 \\ 0, & others \end{cases} \tag{4}$$

   Step three: Determine the next stop. The selection mechanisms to determine the next site can be expressed as:

$$B(i,j) = \frac{p B_p(i,j) + (1-p) B_c(i,j)}{\sum_{j \in v(i)} \left[p B_p(i,j) + (1-p) B_c(i,j) \right]} \tag{5}$$

   Combined with the judging principle of line trend, select $\beta(i, j) * \beta(i, j)$ maximum stop K_j as the next stop of the public transport route. If K_j has been included in the route stop collection N_t, then select the stop which has the second biggest value, and move on.

   Step four: Evolution of line entitlement. Add the new line e_{ti} to the set of the line E_t.

Step five: Volume changes caused by stop connection.

Step six: Judging line terminal. If the line arrives at the terminal station, turn to Step (3); otherwise return to the second step.

3. Complete the evolution of the urban public transport network: if $t = T$, the algorithm is over; otherwise $t = t + 1$, turn to Step (2).

Simulation Analysis

In this paper, the Matlab numerical simulation method is used to study the evolution characteristics of the urban public transport station network with different degrees of the two constraints. First, input the relevant parameters of the network initialization. Assuming that the public transit station network includes 500 stations and 100 routes. The number of stations in each bus route m obey Poisson distribution (=24). The traffic demand $w_{ij}(0)$ obey the power law distribution: $p(w(0)) \sim 4.3 * 10^{-4} * (1 - w(0))^{-2.5}$, $w_{ij}(0)$ is a random number, and $w_{ij}(0) \in (0,1]$. The network edge weight w_{ij} is in accord with the power law distribution: $p(w) \sim 4.3 * 10^{-4} * (1 - w)^{-2.5}$, additional flow $\delta \sim N(0.05, 0.05)$. Secondly, draw the simulation network diagram of evolution model. At last, analyze the basic statistical characteristics of the analog network. Get the intensity distribution map and the line entitlement distribution map, as shown in Figs. 2 and 3.
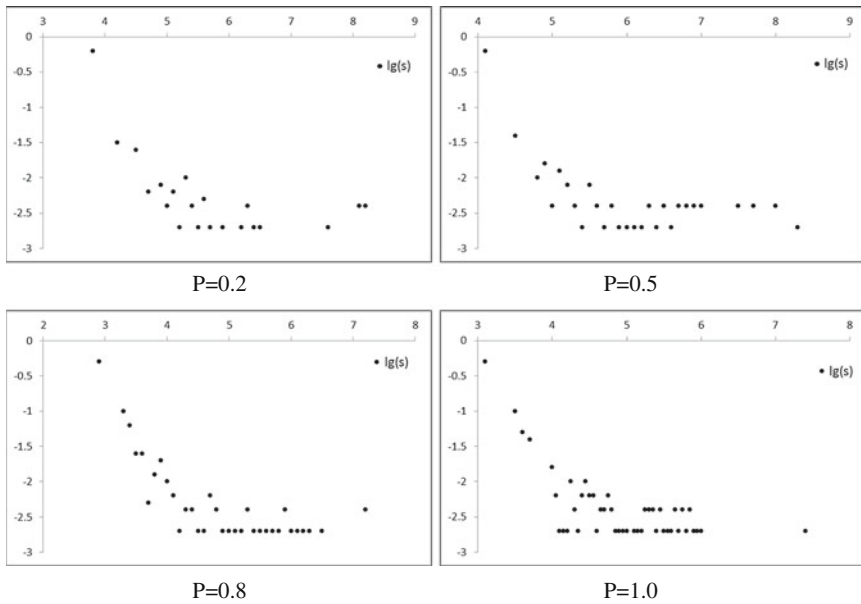
Fig. 2 Public transit station model different interests' mechanism evolution intensity distribution diagram of simulation results
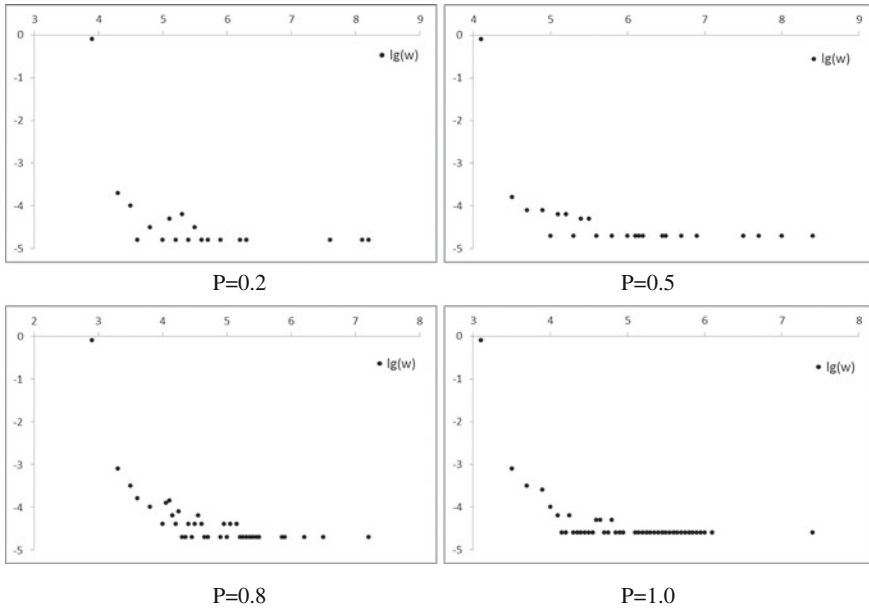
Fig. 3 Different interests under the mechanism of bus station network model of the line entitlement distribution diagram of evolution results

In Figs. 2 and 3, the intensity distribution and line entitlement distribution of the urban public traffic station network are obtained by using the proposed evolutionary model under the different degrees of the two constraints. Double logarithmic coordinate system is adopted in the figure.

From Fig. 1, we can see that with the representation of the passengers benefit restraint mechanism of P value increases, and the intensity distribution of non-scaling properties is more and more obvious, especially at $P \geq 0.8$ to simulate the intensity of urban bus stop network distribution that is almost a straight line and presents the "long tail" phenomenon.

As seen from Fig. 2, with P value increases gradually, the line entitlement of the distribution of non-scaling properties is more and more obvious, especially when $P \geq 0.8$ to simulate urban public traffic station network of the line entitlement distribution is almost a straight line and presents the "long tail" phenomenon, namely network has large amount of small entitlement. And this phenomenon shows that only a small number of sites have a large traffic flow in the simulation of the urban public transport network.

Above analysis shows that: The Space L Mode to construct the urban public traffic station network satisfies with the requirements of the basic traffic flow to and non-linear rate of constraint, and the evolution of public traffic station network should consider passengers benefit restraint mechanism at first, and at the same time to the appropriate balance of investment and operating cost constraint mechanism.

Optimization and Adjustment of Model

Evolutionary Model Construction

On the optimization of public traffic line, it is usually adjusted in the partial urban road network. And this adjustment is affected by many factors, such as bus station does not match the point of passenger source along the public transport route. In a section of the bus line, passengers are divided into two broad categories: (1) The passenger whose destination is not in this segment named passing passenger flow. (2) The passenger whose destination is in this segment named arrival passenger flow. Therefore, when we optimize or adjust the model of public transport network, we should evaluate both the effect of the pass passenger flow and the arrival passenger flow. So, the evolution model has the following three characteristics:

1. As a whole to consider the combined action of the passing and the arriving flow in optimization and adjustment mechanism.
2. The interests of passing flow embodied in travel distance being as short as possible.
3. The optimization adjustment of the network also should consider the interests of the arrival flow. Embodied in network line range is wide and has good accessibility.

Evolutionary Mechanism

The model optimization adjustment mechanisms include:

1. Route growth mechanism.
2. In the restriction mechanism of passing passenger flow: Always aim for segment travel distance being as short as possible (d_{ij}); and it can be defined as

$$B_p(i,j) = \frac{1/d_{ij}}{\sum_{j \in v(i)} 1/d_{ij}} \tag{6}$$

3. In the restriction mechanisms of arrival passenger flow, passengers always want the bus station to have the distance from their own travel destination to be as close as possible, in order to decrease the cost of travel time after they get off. As a result, this will require the routes having better accessibility and larger coverage area. But the bus routes are prone to choose those locations with more traffic demands. Therefore, the restriction mechanism of arrival passenger flow can be defined as

$$B_a(i,j) = \frac{w_{ij}}{\sum_{j\in v(i)} w_{ij}} \qquad (7)$$

Model Evolutionary Processes and Algorithms

Step 1 Initialize network, the network constructed by the Space L Mode.
Step 2 Optimize and adjust the network: where set of nodes is $Nt = \Phi$; set of lines is $Et = \Phi$; and set of line entitlements is $Wt = \Phi$; the optimization plan is generated by using following rules:

1. Determine the origin and destination in adjusting areas in which starting stop and ending stop: Select the origin and destination from the set of local area network nodes. Let $N = \text{N}-\text{O}-\text{D}$;
2. Building the next bus stop. It is still using the line to determine the principle of Sui;
3. Determine the next stop. The selection mechanism to determine the next stop can be expressed as

$$B(i,j) = \frac{pB_p(i,j)+(1-p)B_a(i,j)}{\sum_{j\in v(i)}\left[pB_p(i,j)+(1-p)B_a(i,j)\right]} \qquad (8)$$

4. Judge the stop whether it is the ending stop, if it is the ending stop, go to Step (3); or else return to Step (2)
   Step 3: End of the algorithm.

Simulation Analysis

We use numerical simulation method to study the characteristics of optimal adjustment of the public transit station network in two different degrees of constraint mechanism. In this paper, we select the part of the section of the No. 10 and the No. 24 routes in the public transport network in Guilin as an example, to establish the local bus station network. When carrying out the simulation analysis, stop location and stop OD are simplified process in this paper. We assume that all bus stations in the network location are known, as shown in Fig. 4. And the demand network of the stop is also known. After the location of the stop was known, the matrix of judging direction can be calculated.
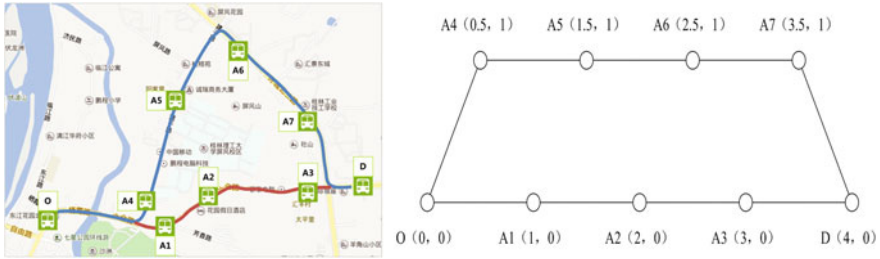
Fig. 4 Local bus station network space location map (*left* actual map, *right* abstract map)

Let $P = 0$, $P = 0.25$, $P = 0.5$, $P = 0.75$, and $P = 1$, get the corresponding local bus station network optimal probability matrix.

Then, it studies the importance of the local optimization and adjustment of the bus station network under the constraint mechanism of passing passenger flow and the mechanism of the arrival passenger flow.

When $P = 0$, the optimized adjustment lines are: O → A4 → A5 → A6 → A7 → D.
When $P = 0.25$, the optimized adjustment lines are: O → A4 → A5 → A6 → A7 → D.
When $P = 0.5$, the optimized adjustment lines are: O → A4 → A5 → A6 → A7 → D.
When $P = 0.75$, the optimized adjustment lines are: O → A4 → A5 → A6 → A7 → D.
When $P = 1$, the optimized adjustment lines are: O → A1 → A2 → A3 → D.

Visible: When P from 0 gradually adjusted to 1, influenced by passing passenger flow and constraint mechanism becomes stronger gradually, the restraint mechanism of the arrival passenger flow gradually becomes weak, the resulting optimization and adjustment of bus lines gradually from the larger flow of station transition to distance of shorter station, so the optimization and adjustment of bus lines gradually from O → A4 → A5 → A6 → A7 → D change to O → A1 → A2 → A3 → D. And, when $P = 0.5$, the optimized adjustment of the bus line is O → A4 → A5 → A6 → A7 → D, reflects the optimization and adjustment of the city bus station network is arrival passenger flow constraints mechanism as the primary consideration. At the same time, when we adjust the urban public transport network, the arrival passenger flow constraint mechanism is the primary consideration.

Conclusions

Based on the complexity analysis of the urban public transport network evolution, this paper determines the urban public transport network which is an entitled and mixed evolution mechanism network. According to the construction of the urban

transit network evolution model and the evolution of the urban public transit network optimization and adjustment model and the simulation analysis of the model can be seen:

1. In the urban public transport network, the optimization and adjustment of the local structure constitute the global network evolution;
2. Whether the local optimization and adjustment, or the overall network evolution, the passenger demands is the primary consideration;
3. In the local optimization and adjustment, the primary consideration is embodied by the arrival passenger flow constraint mechanism. In the global network evolution, the primary consideration is the spatial and temporal constraints;
4. Whether it is local optimization and adjustment of the mechanism of passenger flow constraints or global network evolution of the public transport company cost constraint mechanism, they play the same role on the model changes.

References

1. Wang, Xiaofan, Xiang Li, and Guanrong Chen. 2006. *Complex network theory and application*. Beijing: Tsinghua University Press (in Chinese).
2. Wang, Tao, and Le Xu. 2009. Research on topological structure and evolution model of urban transit network. *Journal of Highway and Transportation Research and Development* 26 (11): 109–112. (in Chinese).
3. Yook, S.H., Jeong, H., and A.L Barabasi, et al. 2001. Weighted evolving networks. *Physical Review Letters* 86(25): 5835–5838.
4. Li, Ying, Wei Zhou, and Shijin Guo. 2007. An analysis of complexity of public transportation network in Shanghai. *Systems Engineering* 5(1): 38–41 (in Chinese).
5. Yu, Bin, Yongzhi Yang, and Zhongzen Yang. 2009. Transit network optimization based on direct Passenger flow density maximization. *Journal of Harbin Institute of Technology* 41 (2): 205–207. (in Chinese).
6. Horner, M.W., and M.E. O'Kelly. 2001. Embedding economies of scale concepts for hub network design. *Journal of Transport Geography* 9 (4): 255–265.
7. Han, Yin, and XiaoGuang Yang. 2005. A new nonlinear bi-level optimization model of transit network and its solution algorithm. *Computer and Communications* 23 (4): 11–14. (in Chinese).
8. Ren, Hualing, and Ziyou Gao. 2005. Research on bi-level model and solution algorithm for dynamic transit design problem. *System Engineering Theory and Practice* 23(4): 11–14 (in Chinese).
9. Dai, Shuai, Huapu Lu, and Qizhou Hu. 2011. *Study on the network optimization based on integrated multi-hiberarchy transit* (in Chinese).
10. Sui, Y., F.J. Shao, R.C. Sun, et al. 2012. Space evolution model and empirical analysis of an urban public transport network. *Physica A: Statistical Mechanics and its Applications* 391 (14): 3708–3717.

A Classification Method for Accesses on Suburban Highway

Yuanyuan Hong, Yongfeng Ma and Jian Lu

Abstract To increase the capacity and to improve the safety on suburban highway, this paper proposes a classification method for access on suburban highway based on the roadside land use type. Considering the standards in China and the field survey, three access categories are defined as follows: Category 1: large residential district or business district; Category 2: hybrid district of commerce and residence; and Category 3: shops along street. The rationality of classification is verified by significance test, which shows that the three categories exhibit significant difference and affect the traffic flow on main road traffic in different degrees. Therefore, the proposed classification method can be used for the basis of relevant researches on suburban highway accesses.

Keywords Suburban highway · Access category · Traffic safety

Introduction

Suburban highway is a product of the economic development and urban expansion, which has the land characteristics of both urban and rural. Nowadays, traffic condition in the zone shows more and more safety problems due to the unreasonable design, the incomplete management, and so on [1]. Therefore, accident rate on suburban highway is relatively higher than other highway segments, especially in the vicinity of access. As a consequence, the access is the key to improving the traffic safety and ensuring the road unimpeded on suburban highway.

Y. Hong · Y. Ma · J. Lu (✉)
Jiangsu Key Laboratory of Urban ITS, Southeast University,
Si Pai Lou No. 2, Nanjing 210096, China
e-mail: lujian_1972@seu.edu.cn

Y. Hong · Y. Ma · J. Lu
Jiangsu Province Collaborative Innovation Center of Modern
Urban Traffic Technologies, Southeast University, Si Pai Lou No. 2,
Nanjing 210096, China

From the 1990s, the USA has carried out research on access systems from whole aspects, including planning, development, design, implementation, and management. And in 2003, the Transportation Research Board Committee on Access Management promulgated Access Management Manual (second edition published in June 2014) [2]. In the manual, the basic roadway functional categories—arterials, collectors, and local roads—are modified and expanded for different planning purposes. Access categories should be internally consistent with roadway functional classifications and with agency transportation plans. Four factors are mainly considered in defining access categories: level of importance of roadways within the jurisdiction, characteristic of system roadways, land use and growth management goals, and current and potential future presence of pedestrians, bicyclists, and transit.

The research on highway access in China started much later than that in the USA, and the existing specification is not perfect. Chao and Jian [3] divided the highway accesses into three levels based on the roadside conflict level. Kunhua et al. [4] focused the analysis on the residents' trip purpose who living around the access in first-class highway and then put forward two types based on the access demands. The two types are survivable traffic demand and living traffic demand— the former mainly grows out of densely populated towns, where generates more demand in long-distance daily trips; the latter includes the shopping, culture, education, and health care, which has a higher request in travel convenience and comfort. Tangzhi et al. [5]. analyzed the access on secondary roads in mountainous area. Four technical indicators (the grade of intersection roads, the passing vehicle models, traffic volume, and the width of the access) were chosen to be the basis of classification. Accordingly, the access on secondary roads in mountainous area was divided into three categories.

Access classification is the foundation of access management. The objective of this paper is to propose a reasonable classification method for accesses on suburban highway based on sufficient spot investigation and correlational research achievements.

Methodology

Preliminary Classification

According to the code for classification of urban land use and planning standards of development land [6], the paper summarizes the land use types on suburban highway in Table 1.

Based on the comprehensive consideration of standard above and the field survey results, three main access categories are put forward and their characteristics are defined explicitly (as shown in Table 2).

Table 1 Land use types on the side of suburban highway

Codes	Roadside land use types	Annotation
R2	Residential land II	Having multistory residence with well-equipped facilities, complete layout, and favorable environment
R3	Residential land III	Having rough residence with insufficient facilities and abominable environment, including dangerous building, shanty town, emergency dwelling, and so on
A1	Administration	Having many organizations, such as party and government offices, social organization, and public institution
B1	Commercial facilities	Providing various commercial service
B4	Public facilities	Providing public service, including refueling, posting, and telecommunications
M1	Industry I	Having industrial enterprise with extremely low interference, pollution, and safety problems to living and public environment
E2	Agriculture and forestry	Having cultivated land, garden land, forest land, grassland, agricultural facilities, Tian Kan, rural roads, etc.

Table 2 Categories and characteristics of access on suburban highway

Categories	Roadside land use types	Feature description
1	Large residential district or business district	Wide access (12–18 m) Vehicle types are mainly cars, vans, and non-motor vehicles Obvious interference in the traffic stream on main road High traffic volume A great many traffic conflict
2	Hybrid district of commerce and residence	Moderate width (8–12 m) Vehicle types are mainly cars, vans, non-motor vehicles with medium-sized trucks, and large trucks occasionally Frequently interference in the traffic stream on main road Moderate traffic volume Passengers and non-motor vehicles have a great impact on lateral path vehicles
3	Shops along street	Small width (4–8 m) Vehicle types are mainly cars, vans, non-motor vehicles with vans occasionally Frequently interference in the major stream Moderate traffic volume Many traffic conflict

Soundness Verification

Evaluation Index

The traffic conflict technology (TCT) originally developed by Perkins and Harris in 1969 as a formal procedure to record near-accidents was used in the safety assessment as the most common compensatory approach [7]. The technique itself is grounded in the ability to register the occurrence of near-accidents directly in real-time traffic and therefore offers a faster and, in many respects, more representative way of estimating expected accident frequency and accident outcomes.

The conflict rate (CR) is selected to be the evaluation index to reflect the safety level of access on suburban highway, which is calculated by the ratio of traffic conflicts (TC) to mixed passenger car unit (MPCU), as shown in Formula 1.

$$CR = TC/MPCU \tag{1}$$

T-Test

T-test is a statistical hypothesis test in which the test statistic follows a Student's *t*-distribution if the null hypothesis is supported. It can be used to determine whether two sets of data are significantly different from each other and are most commonly applied when the test statistic would follow a normal distribution if the value of a scaling term in the test statistic was known.

In this paper, the *T*-test is used to test the differences between the conflict rate data of three access categories. The classification could be reasonable only if the differences between any two categories were significant. The null hypothesis presented below states that two means are equal:

$$\begin{aligned} H_0 &: \mu_1 - \mu_2 = 0 \\ H_a &: \mu_1 - \mu_2 \neq 0 \end{aligned} \tag{2}$$

where

$H_0(H_a)$ the null hypothesis (the alternative hypothesis);
$\mu_1(\mu_2)$ the conflict rate mean of any two categories; and
α the significant level, 95%.

The conflict rate can be tested by an independent two-sample *t*-test, and the *t* statistic is calculated as follows:

$$t = \frac{\overline{X}_1 - \overline{X}_2}{\sqrt{\frac{s_1^2}{n_1} + \frac{s_2^2}{n_2}}} \tag{3}$$

where

$\overline{X}_1\left(\overline{X}_2\right)$	the sample average;
$s_1^2\left(s_2^2\right)$	the unbiased estimator of the variance of any two samples; and
$n_1\left(n_2\right)$	the size of samples.

Furthermore, software SPSS is applied in the research to complete the progressing of data and then analyzes the significance test of classification.

Data Collection

Investigation Method

The investigative tools are an aerial photograph aircraft named Phantom 2 and an onboard Gopro camera produced by SZ DJI Technology Co., Ltd. In the field survey, the aircraft is hovered 200–300 m over the target sections, and the video recorded by the aircraft covers 400–500 m road. A typical scene is illustrated in Fig. 1. The majority of required data can be collected from the videos, including the vehicle composition, traffic flow, and traffic conflict. Furthermore, researchers need to collect the land use condition, highway geometrics, and access information.

Essential Data

One important principle to select the sample access is that no restrictive median was deployed in highway segment. The information about selected seven sample accesses is summarized in Table 3.

The investigation is conducted at the peak hour, and the data of TC and MPCU can be derived from the collected videos. Collision distance (the distance between

Fig. 1 A typical scene of field survey

Table 3 Information of the sample accesses and highway segments

Access number	Access categories	Number of lanes	Median (Y/N)	Limited speed (km/h)
1	1	4	Y	60
2	1	4	Y	60
3	1	4	Y	60
4	2	4	Y	60
5	2	4	N	60
6	3	4	N	60
7	3	4	N	60

car and possible conflict) is selected as an index to judge the accident severity, and when the collision distance is less than the braking distance, we define it as a serious conflict and count it into TC. So the braking distance is the critical distance, which is influenced by different vehicle types and vehicle speed [8]. The index MPCU can be calculated by the traffic flow counting from the videos and equivalence factors referring to Technical Standard of Highway Engineering [9]. And to improve the accuracy of the data, we carried out several data statistics in each sample access. Ultimately, two essential traffic data, TC and MPCU, are summarized in Table 4.

Data Analysis

Conflict Rate

Based on the essential data and Formula 1, the conflict rate (CR) of three access categories is tabulated in Table 4.

With the statistical result shown in Table 4, the access categories can be sorted by the size of CR, Category 1 > Category 3 > Category 2. So the Access Category 1 has the most serious safety problem and exerts large impact on the main road traffic, and the percentage of small vehicle in there reaches up to 71%. Moreover, the Access Category 2 is relatively safer than other access, due to the lowest traffic conflict and conflict rate. And the Access Category 3 shows highest mixed traffic flow and medium level of traffic safety, in which the proportion of small vehicle (only 60%) appears a downward trend, while the proportion of medium vehicle increases significantly to 11%.

Table 4 Essential traffic data and CR of all sample accesses

Category 1				Category 2				Category 3			
No.	TC	MPCU	CR	No.	TC	MPCU	CR	No.	TC	MPCU	CR
1	90	1786	0.0504	4	37	1303	0.0284	6	64	1793	0.0357
	92	1477	0.0623		54	1672	0.0323		79	1876	0.0421
	74	1733	0.0427		24	1538	0.0156		66	1774	0.0372
	75	1648	0.0455		42	1673	0.0251		54	1753	0.0308
	78	1480	0.0527		39	903	0.0432		60	1858	0.0323
	68	1521	0.0447		17	1619	0.0105		36	1614	0.0223
	52	1417	0.0367		27	1698	0.0159		56	1728	0.0324
	96	1512	0.0635		38	1577	0.0241		75	1656	0.0453
	52	1281	0.0406		49	1815	0.0270		27	1534	0.0176
2	32	1265	0.0253		44	1803	0.0244		56	1806	0.0310
	77	1467	0.0525	5	32	1749	0.0183	7	52	1503	0.0346
	90	1705	0.0528		50	1351	0.0370		42	1516	0.0277
	74	1383	0.0535		32	1798	0.0178		56	1573	0.0356
	72	1440	0.0500		42	1667	0.0252		85	1786	0.0476
	76	1479	0.0514		40	1688	0.0237		52	1733	0.0300
3	79	1341	0.0589		35	1535	0.0228		33	1823	0.0181
	67	1232	0.0544		54	1641	0.0329		60	1765	0.0340
	69	1164	0.0593		37	1713	0.0216		48	1543	0.0311
	54	1463	0.0369		48	1412	0.0340		48	1696	0.0283
	66	1535	0.0430		67	1614	0.0415		53	1611	0.0329
AVG	72	1466	0.0489	AVG	40	1588	0.0261	AVG	55	1697	0.0323

Significance Test of Classification

To conduct a significance test, an important prerequisite is the experimental data exhibiting normal distribution. Based on the result of normal distribution examination, the values of skewness and kurtosis are lower than 1.0, and the Asymp.sig. values in KS test are higher than the critical value of 0.5. Therefore, three sample data obey the normal distribution and meet the requirements of T-test.

Then, the ANOVA is taken to verify the difference of the data, and the results of Sig. are lower than 0.05. So the hypothesis H_0 that three groups of data are equal is false under the confidence interval of 95%.

At last, conduct the significance test between any two groups. Three combinations can be listed (Category 1 and Category 2, Category 1 and Category 3, and Category 2 and Category 3). The final results of significance test are shown in Table 5.

Consulting the t critical values, $t_{0.05}(38) = 2.024$, the absolute value of t is higher than 2.024 and the value of sig. is lower than 0.05 in Table 5 which illustrate that there are significant differences between any two access categories. In conclusion, the preliminary classification of access on suburban highway is reasonable.

Table 5 Results of the significance test of classification

	Category 1 & 2	Category 1 & 3	Category 2 & 3
t	7.903	5.999	−2.416
df	37.679	37.905	37.432
Sig. (bilateral)	0.000	0.000	0.021
95% confidence interval	(0.0169, 0.0286]	(0.0116, 0.0235]	(−0.0116, −0.001]

Conclusion

According to the result of field research and theoretical certification, this paper classifies the accesses on suburban highway by the land use types and confirms three main categories ultimately. Access Category 1 is the access in large residential district or business district; Access Category 2 is the access in hybrid district of commerce and residence; and Access Category 3 is the access in shops along street.

Research on suburban highway has an important value for future development and a wide application prospect in China. The research result can be used for the foundation of other correlational research. The further research can pay more attention on influencing factors and collect more sufficient data to subdivide the accesses into several subclasses.

References

1. Yujuan, L. 2011. *Study of alignment optimization technique for urban-rural fringe road based on traffic safety*. Wuhan Institute of Technology.
2. Transportation Research Board. 2015. *Access management manual*, 2nd ed. Washington DC: The National Academies Press.
3. Chao, L., and R. Jian. 2004. *The influence of the entrance and exit density in first-class highway on traffic safety.*
4. Kunhua, T., C. Minglei, T. Jiang, et al. 2011. Study of access on first-class highway. *Highways & Automotive Applications* 03: 25–28.
5. Tangzhi, L., M. Qiang, and M. Zijun. 2013. Safe disposal of countermeasures research on access for mountain distribution of secondary road. *Highway Engineering* 06: 123–127.
6. Code for classification of urban land use and planning standards of development land. 2011.
7. Yan, M., W. Wuhong, and G. Weiwei. 2012. Analysis of opposing left-turn conflicts based on traffic conflict technology. *Journal of Beijing Institute of Technology* 004 (21): 487–491. (English Edition).
8. Yuanyuan, Z. 2006. *Study on fuzzy road safety evaluation of highway intersection based on traffic conflict*. Harbin Institute of Technology.
9. Technical Standard of Highway Engineering. 2014.

A New Method of Code Generation for MC9S12 ECU

Rongge Meng and Chunhua Zhang

Abstract RTW design method on the basis of model is becoming a main method of embedded system development as the short development cycle, low cost, high reliability, ease testing, and maintenance. But during the process of model building, the way of MATLAB/GUI is always used to configure the system parameters, which results in many imperfections such as slow implementation starting, poor compatibility, complex configuration, in-aesthetic interface, and so on. The purpose of this paper is to improve the reliability and accuracy of parameters setting and simplify the configure process by changing the configuring method and simplifying setting process. The parameters of RTW are configured by upper software. The user interface was developed by Visual C# on the basis of three-tier structure. By PWM module testing, the correctness of C# parameters setting, Simulink modeling, and the RTW code generating was verified. The ECU actual running results show that configuring the model parameters by upper software can get rid of the constraints of GUI. The target ECU could be configured conveniently according to the needs of developers whether they are familiar with GUI or not. Meanwhile, more energy could be devoted to the design of control algorithms. Last, by function expansion, the upper software could be used for more controllers to realize the automatic code generation.

Keywords V-mode · RTW · Code generation · ECU

Introduction

V-mode not only combines the software realization and validation but also can shorten the development cycle on the premise of the high reliable software quality. It has been widely used in the field of modern automobile electronics systems [1]. Code automatic generation is a bridge between the product design and the verifi-

R. Meng (✉) · C. Zhang
Chang'an University, Xi'an, China
e-mail: mengrongge@163.com

cation, which plays a key role in V-mode. RTW (Real-Time Workshop) is a model-based approach, which makes it possible to convert the diagram models of Simulink into target C-code automatically. Many researches have been conducted regarding to RTW. In references [2, 3], the method of code automatic generation based on RTW was introduced in the Frascati Tokamak Upgrade Feedback Control System running under RTAI-GNU/Linux system. The method of RTW code generation of automotive electronic control system in OSEK/VDX operating system environment was proposed in paper [4]. Design method on the basis of model has achieved good performance as in automatic DSP code generation [5] and battery management [6]. A great number of researches have been conducted on automobile electric control, and so on. They all use MATLAB/GUI to configure system parameters, which are less than ideal, and little research results have been reported at this point concerning the imperfections of GUI. This paper proposes a new method of parameters setting to supply data for RTW. The user interface is developed by Visual C# based on .NET Framework Platform to realize the parameters configuration in the process of system modeling. The main purpose of this method is to deal with the problems of low speed for GUI starting and running, poor compatibility, and others.

This paper introduces the process of Visual C# model parameters setting, Simulink algorithm design, modeling and simulation, and RTW code generation. There are four parts in this article. Building the user interface based on C# is introduced at first. Details of low driver development on the basis of AUTOSAR 4.0 are discussed in the second section. Then, the realization of MATLAB process layer is studied. At last, the code transplanting and design method verification are presented.

Method

V-mode development process is sketched in Fig. 1. Each stage is in the same platform, which is better for the mistake recognition and the design improvement. Automatic code generation is the key for the product design and realization. Target code is generated automatically according to the setting in the process of ECU development in V-mode. But there are some serious shortcomings of GUI parameter configure: (1) Poor compatibility and slow start, don't support backward compatible. Such as in MATLAB 2006a, there are the panel and button group, which are not available in MATLAB 6.5. (2) It is inconvenient for GUI interface to manage and maintain, meanwhile system variables configuration is complicated. (3) There are a limited number of GUI widget especially for the big program, the user interface cannot been edited beautifully enough. (4) It is difficult for GUI to pass the parameters. Therefore, in this paper, C# is used to develop the user interface to configure the target ECU. Moreover, because .NET class library is used and the pointer type does not support in C#, the access of memory address by program is limited as well as the software program is much more haleness. Data

Requirement analysis and offline simulation

Calibration

Rapid Prototyping

HIL

TargetCode

Automatic code generation

Fig. 1 V-mode development

Provide registers and algorithms
that require initialization

Low Driver C Program

Configuration layer C#

Provide function interface

Provide configuration data
such as XML file

Treatment layer Matlab
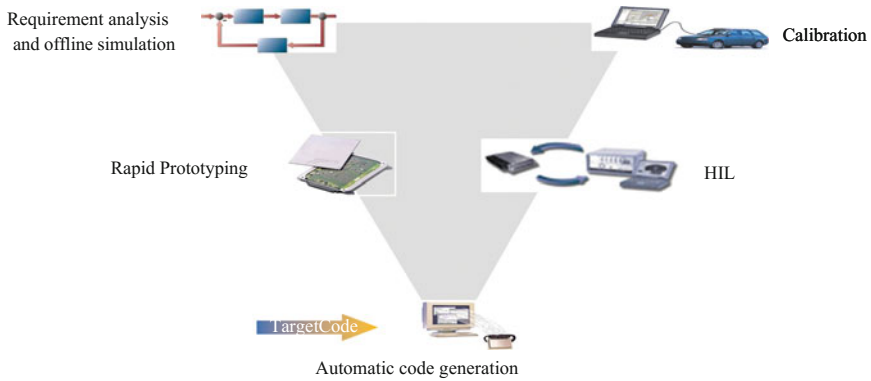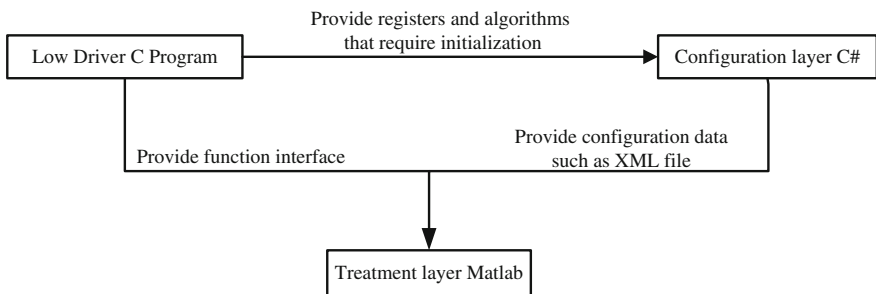
Fig. 2 System code generation architecture

exchange between user interface and Simulink is completed by XML file. The program of lower level driver is written according to AUTOSAR 4.0. Code is generated by RTW toolbox automatically and transplanted to 16-bit platform. The system code generation architecture is realized according to Fig. 2.

C# Interface Development

The software adopted a three-tier structure including presentation layer, business logic, and data access layer from down to up. Presentation layer offered a friendly interface of interacting with computer and displaying the data. The validation of user operation and data exchange was also completed by this layer. The business logic layer between presentation layer and data access was the core part of the system structure; it severed as a connecting link in the process of data exchange. The chief duty of the business logic layer contains the process of critical business, data exchange between presentation layer and data access, complex logic judgment,

and data verification. Database accesses and managements were completed by data access layer. Main purposes of the data access layer in this system were read and written XML file to support data for BLL. The dependence relation among three layers is shown in Fig. 3. In addition, business entity layer which consisted of file and encapsulates properties was built to describe system management objects. Data transmission and management were completed by objects which described by entity class. At the same time, data bonding between presentation layer and business logic was realized. As the data binding among three layers was in one direction, the advantage of high-cohesion and loose coupling for object-oriented programming was produced well. The technologies of data access based on ADO.NET and dynamic data binding were adopted; different dataflows were used to process the different data types. The user operated on interface was related to the data change of dataset and the corresponding XML document. This method guarantees the accuracy and integrity of the data. The application of C# advanced features static polymorphism, interfaces and delegates added the extensibility and maintains of the software. The development cycle was short, cost was low, the user interface was friendly, and the reliability was high.

Compared to the GUI, Visual C# has the following obvious advantages: (1) C# has friendly interface, database operating convenient and data processing easy. (2) Multithreading of C# supports a good way to enhance the parallel task processing ability. (3) GUI needs to install MCR and configure the system variables when it breaks away from MATLAB environment. Instead, .NET framework sets
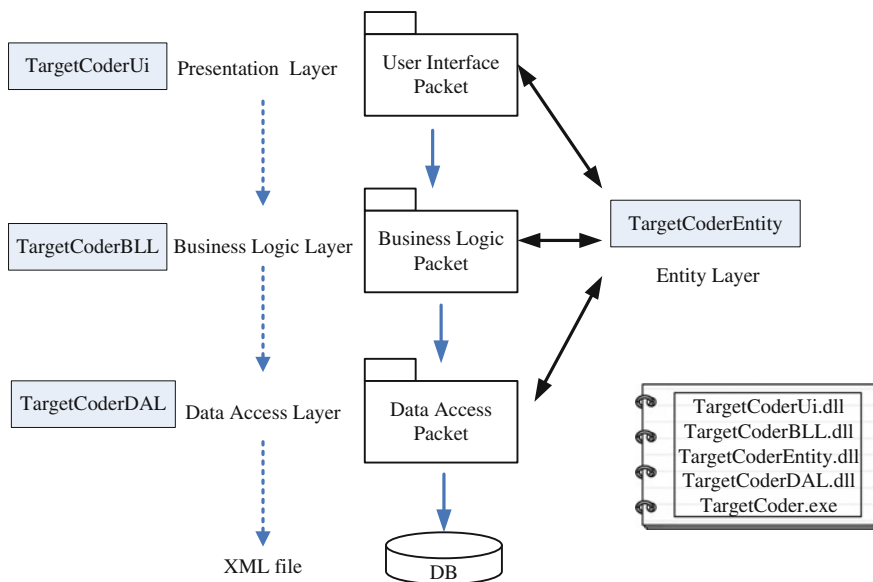
Fig. 3 Three-tier architecture of the user interface
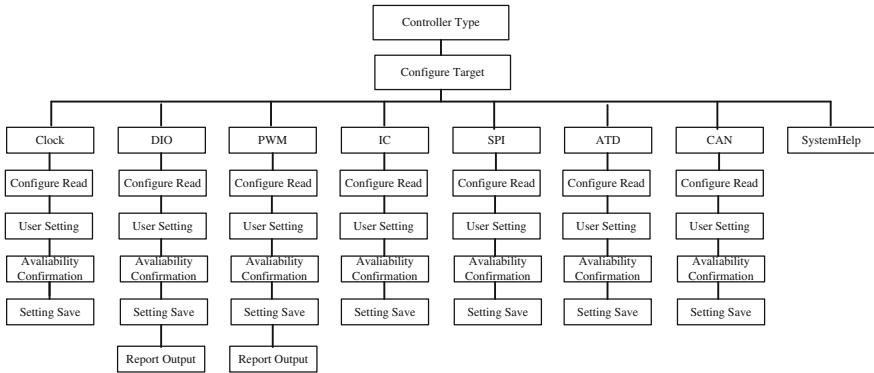
Fig. 4 Visual C# user interface frame

up easily and has a good performance in source code transplant. The software data transmission among three-tier structures has strong dependence and in one direction. Data are transported from expression layer to business logic at first and then transmitted to data access layer from business logic. The translation among three layers was carried out with the assistance of entity layer. Three-layer design has three advantages as follows: (1) The project modification and maintenance become easy. Three-tier structure designs make it easy to modify the function in the process of development, expand and promote the system after completing, and transplant code at last. Simple modification of the corresponding layer does not require to change the other two layers. (2) System function expanding is easy to implement. Expanding is realized just by adding a method in corresponding class, and it is of no need to modify the project structure. (3) It is easy to achieve code reuse and division of labor cooperation. Visual C# software supports a friendly user interface and gets away from the constraints of the MATLAB. Because of the C# interface, users can configure the target board conveniently and generate the code automatically according to their own needs whether they are familiar with GUI or not. Data exchange between Visual C# and Simulink model is completed by XML. Simulink model configures the model parameters and generates the code by reading the user design on the software. Actual application showed that the parameters configured by software made the process of code generation much simpler and more reliable. The Visual C# software frame is indicated in Fig. 4.

Low Driver Development

AUTOSAR is a general embedded system architecture for automotive industry, its core idea is "unified standard, distributed implementation, centralized configuration." Software interface is unified, the system is modularized and hierarchical, and
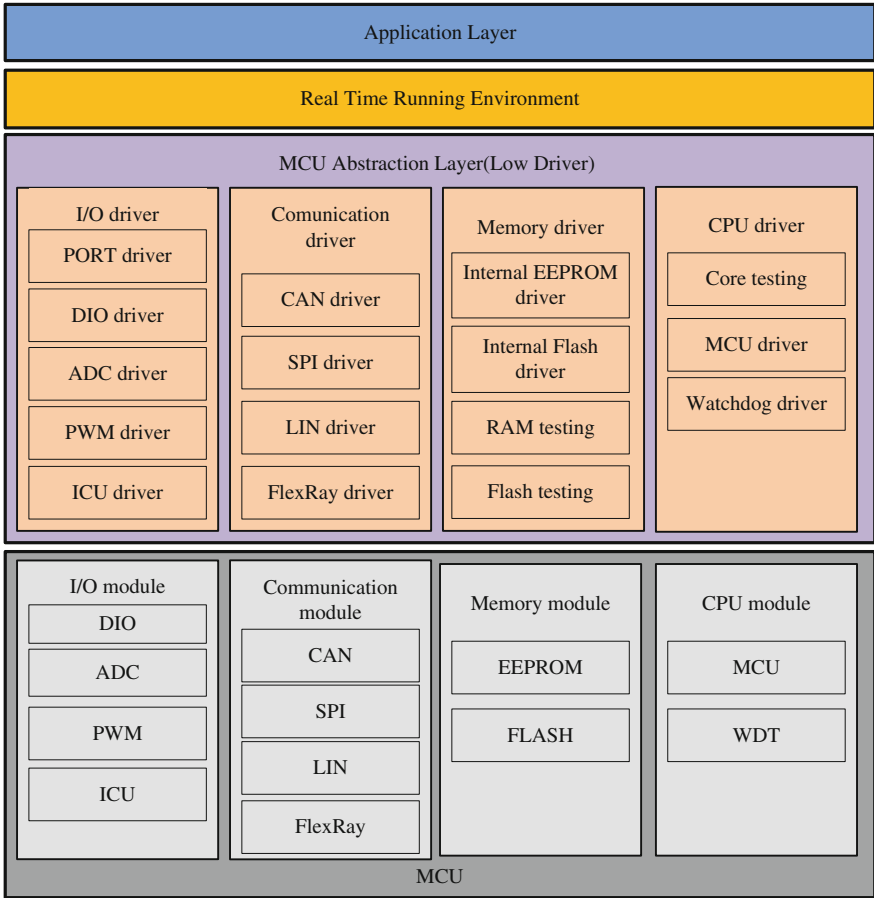
Fig. 5 System low driver framework

module information has centrally unified the format, management, and configuration [7]. On the basis of AUTOSAR architecture standard, the system low driver framework is developed according to Fig. 5.

Process Layer MATLAB

RTW (Real-Time Workshop) is an important toolbox in MATLAB, which provides a code generation environment for users based on Simulink. RTW can convert the Simulink model into transplantable and customizable efficient code. The code generation process is shown in Fig. 6.
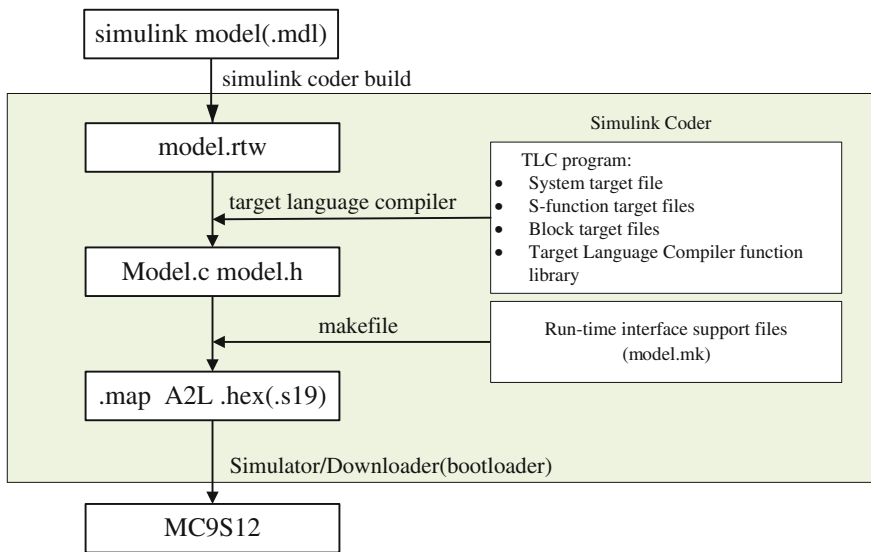
Fig. 6 Code generation process

The Simulink model (model.mdl) was compiled to .rtw file by building the command at first, then the .rtw file was converted to .c source file, and .h head file by target language compiler according to tlc file and library file. Three kinds of tlc files were included, system target file (self-programmed s12.tlc), block target files (.tlc which go with its S-function), and target language compiler function library. The system target file was used to specify the whole structure of generate code to matching the target chip and language. S-function file was used to execute the process of simulation. The model target file was responsible for code generation by obtaining the parameters information from S-function during the model compilation. Meanwhile, the model tlc file had the same name with its S-function. After the model source file and head file generated, by run-time interface support files, the target files (.map, A2L, and .hex) were generated. At last, the code was downloaded to ECU board MCS12.

Results

According to the requirement analysis of ECU development, 16 models were developed at last, shown in Fig. 7. It consisted of ADC, CAN, CCP, DIO, EEPROM, ICU, Init, OCU, PIT, PWM, REST, SPI, Interrupt, BaseClock, Utilities, and MIOS model. These models were saved in user-defined model library in Simulink browser.

Fig. 7 Developed models
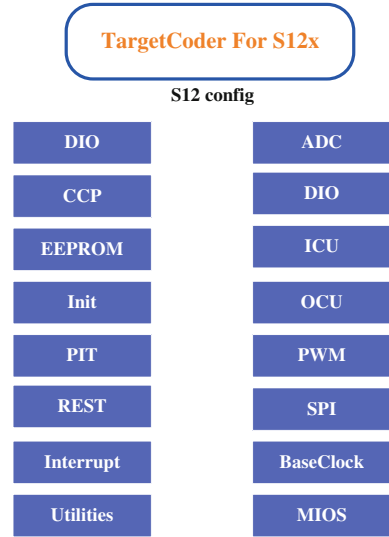
RTW design method on the basis of model has been well accepted to develop the embedded system, as the short development cycle, low cost, high reliability, ease testing, and maintenance. The upper user interface which developed by Visual C# was used to configure the parameters in system modeling. TargetCoder, the upper software development platform based on .NET Framework realized the main parameters configuration for the underlying drivers, including target configuration, clock, digital input and output, pulse width modulation, enhanced capture timer, serial peripheral interface, analog to digital conversion, periodic interrupt, and the module of CAN configure. The operator interface of TargetCoder is shown in Fig. 8.

TargetCoder and RTW data can be exchanged by XML file format. Data communication between upper software and RTW was conducted directly. The testing model PWM was built, shown in Fig. 9, and the generated code is shown in Fig. 10.

The test realized the PWM module code generation automatically by Visual C# parameters setting, Simulink modeling, and RTW code generation. The output frequency was 50 Hz, and the duty cycle was 20%. The generated code is downloaded to VCU1602-0100 ECU, and the connector 111 is corresponding to PIN001 PWM3 of S12. The waveform is observed by oscilloscope TPS2024B, shown in Fig. 11.

The result verifies the correctness of interface setting, modeling, and the generated code. Meanwhile, as the low drive chip was used in PWM channel of ECU, the PWM channel response is slow and PWM frequency could not be too fast.
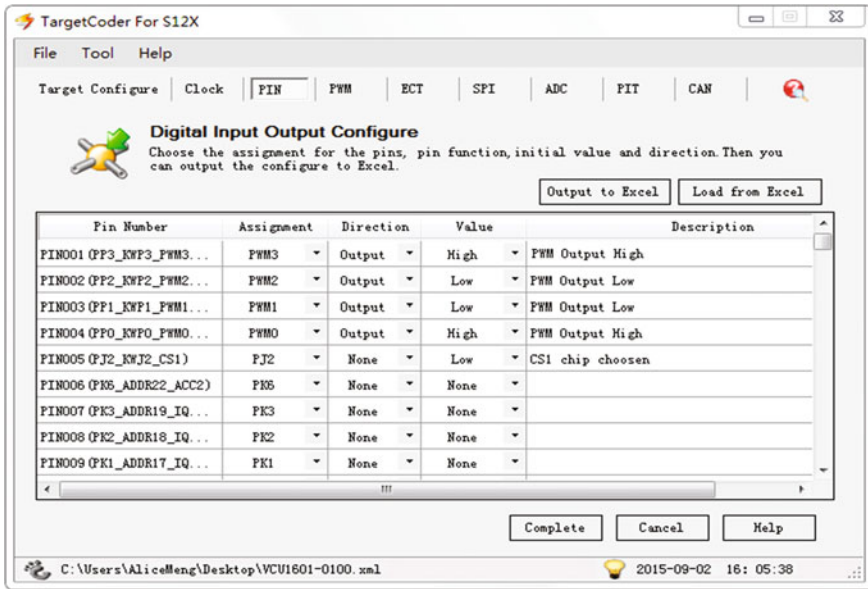
Fig. 8 Developed user interface TargetCoder
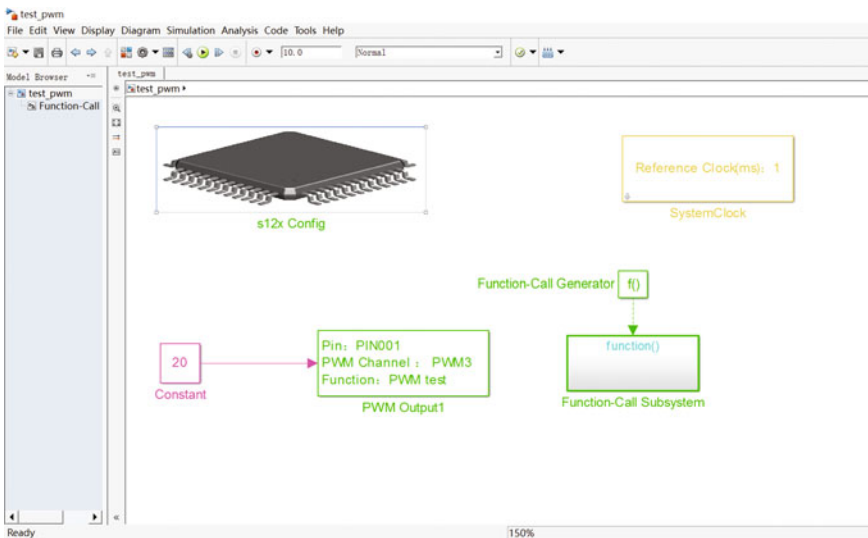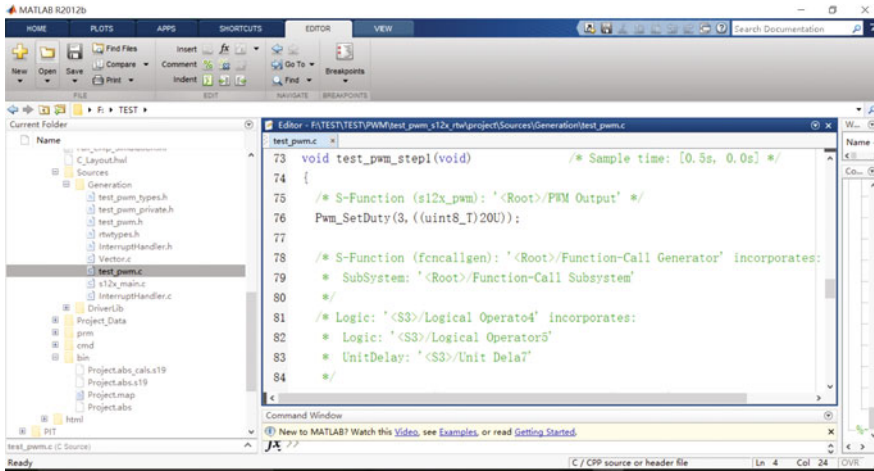
Fig. 9 PWM model testing

Fig. 10 PWM model generated code
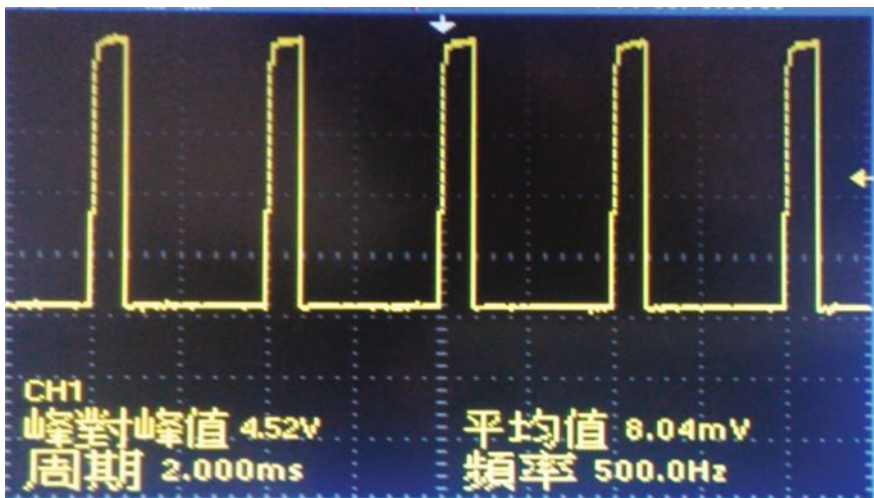
Fig. 11 PWM waveform

Conclusion

The aim of this paper was to optimize the process of RTW code generation auto-
matically by changing the parameters setting method. This paper changes the way
of traditional system parameters configuration that was completed by
MATLAB/GUI directly, and this study explored the use of C# to configure the

parameters. The code generation process showed that the proposed method has the following advantages compared to MATLAB GUI. First, it can simplify the parameters setting especially for some complex interface and algorithm. Second, C# can improve the reliability and accuracy of parameters configuration. Last, the portability of C# code can add the expansibility of the system, and it can be also used on 32-bit ECU after the function expansion. RTW can generate C-code automatically by configuring the parameters simply and building a simulation model. By this way, software design is in isolation from the hardware, the dependence on hardware developer is abated effectively, and the reliability of the program is increased obviously. Therefore, the developer can devote much more energy to the design of control algorithms. So, this method has important practical value.

References

1. Qi, Zhang. 2014. Motor controller design using dSPACE and based on V-Mode development. *Research and Exploration in Laboratory* 33 (11).
2. Vital, V., C. Centioli, F. Iannone, M. Panella, L. Pangione, M. Sabatini, L. Zaccarian, and R. Zuccala. 2005. Using real time workshop for rapid and reliable control implementation in the Frascati Tokamak Upgrade Feedback Control System running under RTAI-GUN/Linux. *Fusion Engineering and Design* 74: 593–597.
3. Vital, V., C. Centioli, F. Iannone, M. Panella, L. Pangione, M. Sabatini, L. Zaccarian, and R. Zuccala. 2007. A Matlab based framework for the real-time environment at FTU. *Fusion Engineering and Design* 82: 1089–1093.
4. Zuo, Wenying, Yingguo Li, Fengjuan Wang, and Xiaobo Hou. 2012. A new design method of automotive electronic real-time control system. *Physics Procedia* 25: 1126–1132.
5. Zhao, Jiqian, and Yaoqin Jia. 2010. The development technology of DSP code based on Matlab/Simulink. *Power Electronics* 44 (12).
6. Xu, Chengjiao. 2014. Automatic code generation based on targetlink and its application in battery management system. *Research on Development* 4.
7. Wei, Xuezhe, Haifeng Dai, and Zechang Sun. 2012. Methodology, architecture and development flow of automotive embedded systems. *Journal of Tongji University (Natural Science)* 40 (7).

Simulation and Evaluation of Guidance Strategies in the Park and Ride Condition

Liu Yang, Wei Guan and Wenyi Zhang

Abstract Traffic guidance strategies have been investigated in the studies of intelligent transportation systems. Most of the prevailing studies are carried out at the single-mode traffic network, while few studies consider the multimodal traffic network, which links with metro service in the Park and Ride (P&R) condition. To fill the gap, this paper builds a multimodal traffic network, which includes both the road and the metro, where the traffic simulator Paramics is conducted to run the simulation. Bang-bang, P, and PI strategies are selected as the guidance strategies, and the average travel cost is utilized as the evaluation indicator. The results indicate that PI strategy has better performance than P and bang-bang strategies in the normal period, while PI and P strategies have similar tends and work better than bang-bang strategy in the peak period. Furthermore, when taking parking spots into consideration, the more the parking spots, the better the guidance performance. These findings compare the performance of different guidance strategies in the P&R condition, and can be applied for transportation systems operations and control, thus to alleviate the urban traffic congestion.

Keywords Traffic guidance strategy · Multimode traffic · Park and ride · Intelligent transportation system · Paramics

L. Yang · W. Guan (✉) · W. Zhang
MOE Key Laboratory for Urban Transportation Complex Systems
Theory and Technology, Beijing Jiaotong University, Beijing, China
e-mail: weig@bjtu.edu.cn

L. Yang
e-mail: summeryangliu@bjtu.edu.cn

W. Zhang
e-mail: wyzhang@bjtu.edu.cn

Introduction

Urban traffic flow guidance system (UTFGS) is one of the subsystems of advanced traffic management systems, which can help drivers to choose the optimum route by analyzing the real-time traffic data [1, 2]. The core content of UTFGS is the guidance strategy, the performance of which decides the guidance effectiveness of the whole traffic network. Hence, it is of great significance to evaluate different guidance strategies using traffic simulation software.

The studies on the guidance strategy started relatively early. Papageorgiou [3] first applied the optimal control method and feedback control framework to the dynamic traffic assignment, and a macroeconomic model framework of a different road network was proposed. Messmer and Papageorgiou [4] built a multi-origin and multi-destination freeway traffic networks. In this paper, a number of classic control strategies were treated as the guidance strategies to adjust the road guidance splitting rates and the effectiveness of which was tested by the simulation. Wang and Papageorgiou [5] took simple highway network as the target network to investigate the effectiveness of bang-bang, P, PI, and LQI controllers and studied various factors' influence of guidance performance by the highway simulation software METANET. Furthermore, the feedback strategies and iterative strategies were compared in a complex highway network and the results indicated that the feedback strategies based on reactive travel time had similarly good performance with the iterative strategies [6]. Kotsialos and Papageorgiou [7] used discrete optimum control to describe highway network problem, which took traffic control strategies design into consideration and testified many different control strategies in a simple highway network. The results showed that the optimal control approach was found to have the efficiency, flexibility, and intelligent control behavior. Xu et al. [8] proposed a feedback optimal control model for urban freeway, which was based on the model-based predictive control framework. The simulation results indicated that this model can explain the time-dependent traffic characteristics and guarantee a high degree of control benefits.

Previous studies usually selected the highway network as the test network, which was built for the evaluation of guidance strategies, while Deflorio [9] chose the urban road network to evaluate the performance of a dynamic route guidance strategy. Florian et al. [10] built an iterative dynamic traffic assignment model, which is based on simulation and applied the model to Stockholm's real road network. The results suggested that this model was applicable to medium-sized networks with a very reasonable computation time. Zhao [11] established a dynamic guidance model based on the real-time network information and red light delay. Then, a numerical experiment was conduct to investigate the feasibility and effectiveness of the algorithm. Xu et al. [12] considered that travelers in traffic network not only had route choice preference, but also had time preference. Hence, the guidance strategies based on either time or route choice preference were proposed and tested to be valid. He and Guan [13] proposed a hybrid strategy, which considered both travel time and traffic state as a solution of conflict between

time-varying demand and long intervals. The simulation results showed that the hybrid strategy can meet the requirement both in the stability and accuracy under long update intervals.

The studies of traffic guidance strategies are various. However, most studies focused on single-mode traffic network, for example, the study object was either the highway network or the urban road network. The multimodal traffic network, which includes both the road network and the metro network in the condition of P&R, was seldom considered. Furthermore, as a kind of traffic management solution to the urban traffic congestion [14], P&R is advocated by the governments and there are comparatively enough supporting facilities in many cities. Hence, the P&R condition is well worth studying. To fill the gap, a metro route with a P&R parking lot is added to the road network and several feedback guidance strategies are evaluated using Paramics simulation software. The simulation experiments compare the performance of bang-bang, P, and PI strategies in normal traffic flow period and peak traffic flow period. Furthermore, the influence of the number of P&R parking spots on the guidance effectiveness is also evaluated.

The rest of this paper is organized as follows: Section "Feedback Traffic Guidance Strategies" introduces the feedback traffic guidance strategies, which includes bang-bang, P, and PI strategies. Section "Multimodal Traffic Travel Cost Model" describes the multimodal traffic travel cost model. Section "Simulation Test" conducts the simulation test and the results analysis. Conclusions for the analysis results are made in Section "Conclusion".

Feedback Traffic Guidance Strategies

Dynamic User Equilibrium

The goals of dynamic traffic guidance can be classified into two types: one is dynamic user equilibrium, and another is dynamic system optimality. In this paper, dynamic user equilibrium is selected as the target of guidance strategies according to the phenomenon that the travelers always concentrate on their own benefits in the real world. When dynamic user equilibrium is achieved, all the routes between a single OD pair have the same travel time, and other routes with longer travel time will not be used.

$\Delta\tau_{n,j}(k)$ means the travel time difference from a bifurcation node n to the destination j, p and s are two segments, where p is the primary route and s is the second route. The travel time difference [15] between p and s is as follows:

$$\Delta\tau_{nj}^{sp}(k) = \tau_{n,j}^{s}(k) - \tau_{n,j}^{p}(k). \tag{1}$$

When there are only two routes in the road network, the smaller the travel time difference $\Delta\tau_{n,j}(k)$, the better the traffic guidance performance. $\Delta\tau_{n,j}(k) = 0$

indicates that the dynamic user equilibrium is achieved. Hence, the travel time difference $\Delta\tau_{n,j}(k)$ is a significant indicator to evaluate and compare the guidance effects of different guidance strategies.

Feedback Guidance Strategies

Bang-Bang Strategy

Bang-bang strategy [6] is quite similar to all-or-not assignment, actually it means when drivers reach the bifurcation node n, all the drivers will be guided to the shortest route between bifurcation node n to destination j, and the formula is as follows:

$$\beta_{n,j}(k) = \begin{cases} 1 & \text{if } \Delta\tau_{n,j}(k) \geq 0 \\ 0 & \text{if } \Delta\tau_{n,j}(k) < 0 \end{cases}. \tag{2}$$

In the formula above, $\beta_{n,j}(k)$ denotes the splitting rate of the (n, j)-couple, namely, the origin–destination couple, during guidance period k; $\Delta\tau_{n,j}(k)$ is the travel time difference of the (n, j)-couple.

The advantage of bang-bang strategy is that the calculation is very simple as well as convenient. Furthermore, it also does not need to set any parameters. The limitation of bang-bang strategy is that it is easy to lead to the oscillation of travel time difference and splitting rate.

P Strategy

P strategy [6] takes the effect of travel time difference $\Delta\tau_{n,j}(k)$ into consideration and introduce adjustable parameter K_p, and its formula is as follows:

$$\beta_{n,j}(k) = \beta_{n,j}^N + K_p\Delta\tau_{n,j}(k), \quad \beta_{n,j}(k) \in [0, 1] \tag{3}$$

where $\beta_{n,j}^N$ denotes the nominal splitting rate describing drivers' behavior of route choice in the absence of route guidance, and typically, $\beta_{n,j}^N$ is set equal to 1, i.e., K_p is the adjustable parameter. The rest of the symbols have the same meaning as former.

It can be known from formula (3), with the increase of adjustable parameter K_p, when K_p exceeds the critical value, the system will come out high oscillation. If parameter K_p is sufficiently high, the P strategy will change into the bang-bang strategy. In addition, when the system reaches the steady state, P strategy still cannot make travel time difference reach zero, in other words, $\Delta\tau_{n,j}(k) \neq 0$.

PI Strategy

PI strategy [6] is based on the former splitting rate and not only considers the influence of present travel time difference, but also the former travel time difference. Two adjustable parameters K_p and K_i are introduced. The PI strategy can be formulated as

$$\beta_{n,j}(k) = \beta_{n,j}(k-1) + K_p[\Delta\tau_{n,j}(k) - \Delta\tau_{n,j}(k-1)] + K_i\Delta\tau_{n,j}(k), \quad \beta_{n,j}(k) \in [0,1] \tag{4}$$

where K_p and K_i are the proportional and integral gains, respectively. $\beta_{n,j}(k-1)$ is the splitting rate of the (n, j)-couple during guidance period $k-1$; $\Delta\tau_{n,j}(k-1)$ is the travel time difference of the (n, j)-couple during guidance period $k-1$. The rest of the symbols have the same meaning as former.

It can be known from formula (4) that with the increase of adjustable parameters K_p, K_i the system will also have oscillating behavior. Comparing to P strategy, the advantage is that PI strategy can make travel time difference reach zero in the steady state, namely, $\beta_{n,j}(k)$ and $\Delta\tau_{n,j}(k)$ will not change with time, independently of the values of the (constant) disturbances (e.g., demand, OD). In addition, another parameter K_i is introduced, and thus it is really difficult for parameter choice.

Multimodal Traffic Travel Cost Model

The traffic network used in this paper is a multimodal traffic network, which includes both the road route and the metro route. In order to compare the road network with the metro network, a multimodal traffic travel cost model [16] is introduced to unify the travel costs in the road network and the metro network.

The Generalized Travel Cost of the Road Network

The car is a kind of private travel tool, although it has a high level of comfort, it also has a high monetary cost, and the travel time is significantly influenced by road congestion. Take all these characteristics into consideration, the generalized travel cost of the road network can be divided into three parts: travel time, monetary cost, and comfortable degree loss. The generalized cost function of car network is as follows:

$$C_a^r = T_a^r + \text{mot} \times \text{mfactor} \times L_a^r + \text{sot} \times \text{sfactor}^r \times T_a^r, \, a \in A_r, \tag{5}$$

where r denotes the road network; a denotes the segment; A_r stands for link set of the road network; T_a^r is the car travel time on segment a; mot presents

monetary-time conversion coefficient; mfactor stands for the unit mileage monetary cost, namely the fuel charge; L_a^r is the length of segment a; sot is the comfort-time conversion coefficient; and sfactor^r presents the vehicle comfortable degree decrease per hour. Furthermore, if the segment involved highway toll, the toll should be added to monetary cost.

The Generalized Travel Cost of the Metro Network

Urban rail transit is a kind of public transit, comparing with traveling by car, the comfortable degree is lower. The metro system has independent road right, thus the metro route will not be influenced by the traffic congestion. Furthermore, the P&R parking lot is built in the network, and thus, the transfer time of car to subway and the waiting time of subway should be considered. Hence, the generalized travel cost of subway network can be divided into three parts: travel time, transfer cost, and comfortable degree loss. In transfer cost, travel time, monetary cost, and transfer expense are included. The generalized cost function of car network is as follows:

$$
\begin{aligned}
C_a^m = T_a^m + \left(T_a^H + \text{mot} \times \text{pay}_a + \text{transfactor} \times T_a^H \right) \\
+ \text{sot} \times \text{sfactor}^m \times T_a^m, \quad a \in A_m
\end{aligned}
\tag{6}
$$

where m denotes the metro network; a denotes the segment; A_m stands for link set of the metro network; T_a^m stands for the travel time on segment a; T_a^H means the transfer cost; mot presents monetary-time conversion coefficient; pay$_a$ stands for the ticket price from origin to destination via segment a; transfactor denotes the transfer penalty factor, which is a positive number without dimension; sot is the comfort-time conversion coefficient; and sfactor^m presents the subway comfortable degree decrease per hour.

According to the multimodal traffic travel cost model above, the generalized travel cost can be calculated. The travel time in the feedback strategies will be replaced by generalized travel cost.

Simulation Test

Description of the Test Network

The P&R parking lots are usually built in the junction between urban area and outskirts. In this paper, Beijing Xierqi subway station, which is the junction of Haidian district and Changping district, is selected as the bifurcation node of object network. The multimodal traffic network is established according to the routes between Beijing Xierqi–Xizhimen in the real world and the Xierqi–Xizhimen network is shown in Fig. 1.
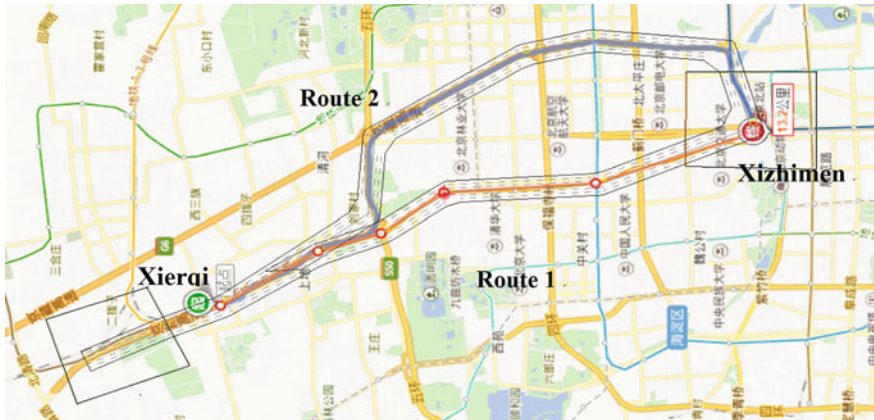
Fig. 1 Xierqi–Xizhimen network model

There are two routes in the Xierqi–Xizhimen network: Route 1 is the subway line 13, the length of Route 1 is about 13.7 km; Route 2 is the shortest road, which includes part of G6 Jingzang Expressway and the length of which is about 17.7 km. The Xierqi subway station has a P&R parking lot, which lies in the southwest of Xierqi subway station. The P&R parking lot has 191 parking spots and the area of which is 7400 m^2. In addition, considering the toll station of G6 Jingzang Expressway, a bottleneck of 5 km/h speed limit is set at the entrance link of the toll station.

Simulation Parameter Design

The simulation experiment has a duration of 1 h for each scenario. The guidance period, traffic demand, and the number of parking spots are selected as the variables. The parameters of these variables are as follows: (1) The guidance interval: 1, 5, and 10 min; (2) The traffic demand: 4600 pcu/h (in the peak period) and 3000 pcu/h (in the normal period); and (3) The number of parking spots ranges from 200 vehicle to 1200 vehicle. Furthermore, in order to focus on the guidance performance of each of the guidance strategies, we only set demand data from Zone 1 (Xierqi) to Zone 2 (Xizhimen).

Simulation Results

The traffic demands and guidance periods are various, and thus, the analysis of guidance strategies is divided into normal period and peak period. The travel cost

difference is treated as the evaluating indicator for the comparison of guidance performance, and the smaller the travel cost difference, the better the guidance performance.

The guidance performances of bang-bang, P, and PI strategies in the normal period are shown in Fig. 2.

As shown in Fig. 2, the PI strategy has the best guidance performance in the normal period, especially when the guidance interval is 5 min. Furthermore, the guidance performance of the P strategy is better than the bang-bang strategy and both the P and PI strategies have similar performance in the 1 min interval and the 10 min interval. In a word, the travel cost difference curves of these guidance strategies do not have large fluctuations and the trends of which are decreasing gradually in normal period. The decreasing trends indicate that the more the parking spots, the better the guidance performance.

The guidance performances of bang-bang, P, and PI strategies in the peak period are shown in Fig. 3.

As shown in Fig. 3, the P and the PI strategies have the similar change trends as well as the guidance performance. Furthermore, both of them have better guidance performance than the bang-bang strategy. When the guidance period is 1 min, the P and the PI strategies have better guidance performance. Thus, the PI and P strategies have better performance in the short guidance interval. In addition, the travel cost difference curves of these guidance strategies are gradually decreasing, and it demonstrates that the parking spots really make a difference in the peak period and the more parking spots, the better guidance performance.

Conclusion

Traffic guidance is an effective method to alleviate the urban traffic congestion and improve the efficiency of transportation management. The guidance strategies play an important role in the guidance effects, and thus it is of great significance to evaluate the performance of various guidance strategies. Most previous studies focus on single-mode traffic network, while the multimodal traffic networks in the P&R condition are rarely considered. Thus, the major contribution of this study is to evaluate the performance of different guidance strategies in a multimodal traffic network and investigate the influence of R&R parking lots. The results indicate that PI and P strategies have better performance than bang-bang strategy in the normal period. In peak period, all these strategies have the similar performance when the guidance intervals are 5 and 10 min, and PI and P strategies have better performance in 1 min guidance interval. Hence, the PI and P strategies are more reliable and stable than the bang-bang strategy. In addition, all the travel cost differences decreases with the increase of P&R parking spots, and thus the number of P&R parking spots relates to the guidance performance and a large number of P&R parking spots will contribute to a better guidance performance.

(a)

(b)

(c)
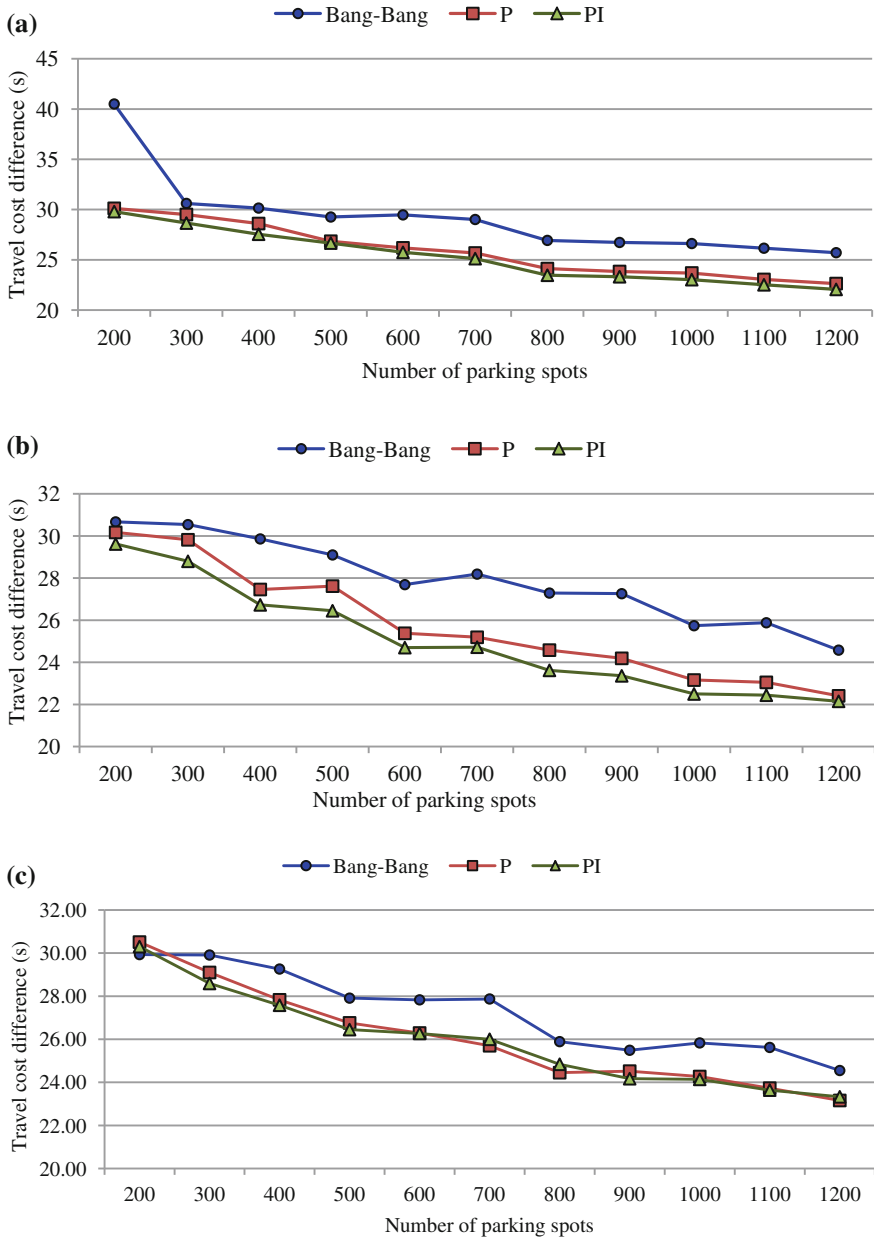
Fig. 2 The guidance performance of bang-bang, P, PI strategies in normal period. **a** Travel cost difference of 1 min period. **b** Travel cost difference of 5 min period. **c** Travel cost difference of 10 min period

(a)

(b)

(c)
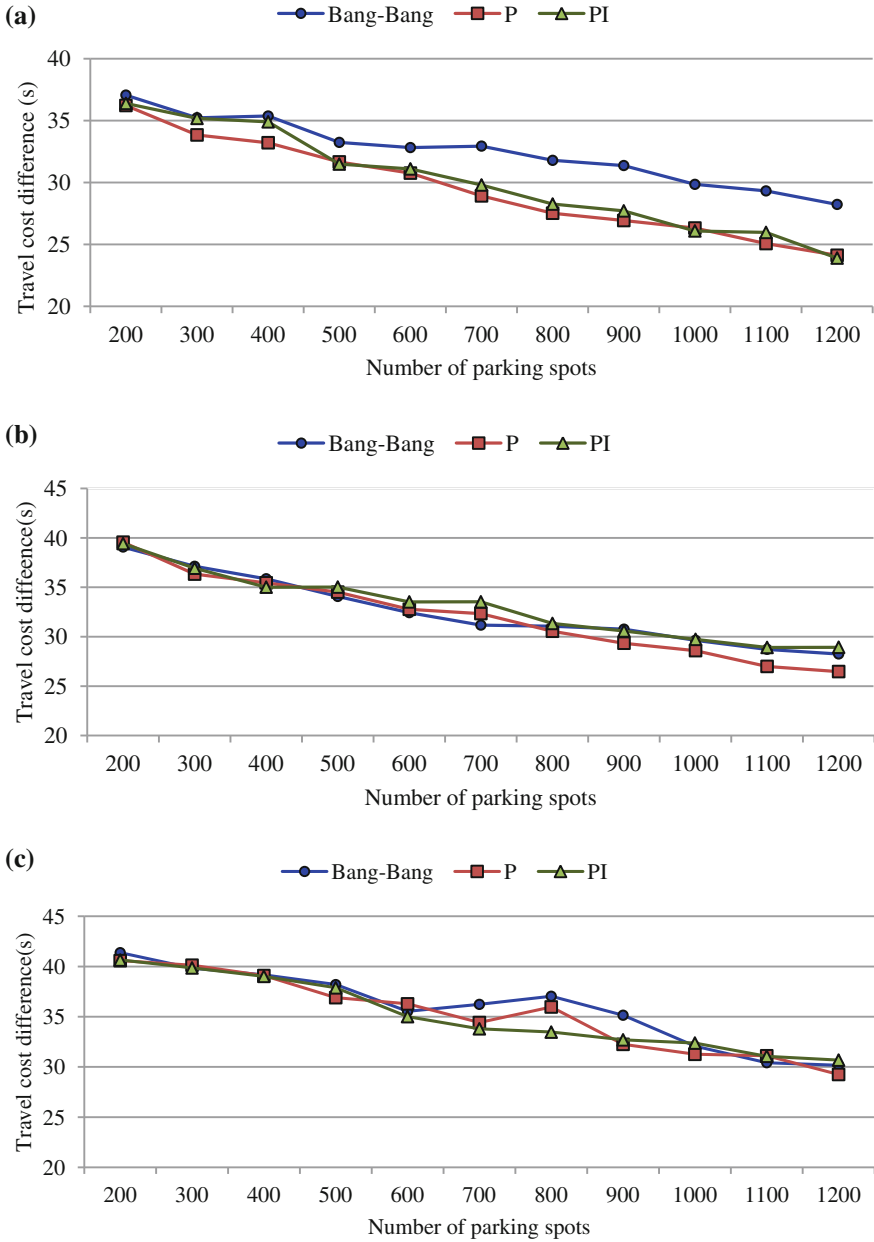
Fig. 3 The guidance performance of bang-bang, P, PI strategies in peak period. **a** Travel cost difference of 1 min period. **b** Travel cost difference of 5 min period. **c** Travel cost difference of 10 min period

On the basis of the performance of different guidance strategies in the multi-modal traffic network, future direction of the work would be to design a new guidance strategy based on multimodal traffic network, which can be treated as a supplement for the current guidance strategies, and help to alleviate the urban traffic congestion. In addition, we would also take the construction cost of P&P parking lots into consideration and investigate optimal scale of P&R parking lots. Thus, the results can be utilized for the design of P&R systems and make this green travel mode more accessible.

References

1. An, S., N. Cui, and C.B. Hu. 2008. Simulation and analysis of route guidance strategy based on a multi-agent-game approach. In *International Conference on Management Science and Engineering*, 140–146. IEEE.
2. Yang, Zhaosheng, Peng Gao, Mo Li, and Nan Zhang. 2011. Applicable prevention method of braess paradox in urban traffic flow guidance system. *International Journal of Computational Intelligence Systems* 4(6): 1254–1261.
3. Papageorgiou, M. 1990. Dynamic modeling, assignment, and route guidance in traffic networks. *Transportation Research Part B: Methodological* 24 (6): 471–495.
4. Messmer, A., and M. Papageorgiou. 1994. Automatic control methods applied to freeway network traffic. *Automatica* 30 (4): 691–702.
5. Wang, Y., and M. Papageorgiou. 2000. Feedback routing control strategies for freeway networks: A comparative study. In *Proceedings of the 2nd International Conference on Traffic and Transportation Studies*, 642–649. ASCE.
6. Wang, Y., M. Papageorgiou, and A. Messmer. 2001. Feedback and iterative routing strategies for freeway networks. *Control Applications* 1776: 1162–1167.
7. Kotsialos, A., and M. Papageorgiou. 2002. Coordinated and integrated control of motorway networks via non-linear optimal control. *Transportation Research Part C: Emerging Technologies* 10: 65–84.
8. Xu, Tiandong, Lijun Sun, and Yuan Hao. 2010. Model-based predictive route guidance method for urban freeway network. *Journal of TONGJI University (Natural Science)* 38(6): 827–838.
9. Deflorio, F.P. 2003. Evaluation of a reactive dynamic route guidance strategy. *Transportation Research Part C: Emerging Technologies* 11 (5): 375–388.
10. Florian, M., M. Mahut, and N. Tremblay. 2004. Application of a simulation-based dynamic traffic assignment model. *European Journal of Operational Research* 189 (3): 1381–1392.
11. Zhao, Huibin. 2011. *Dynamic Guidance Strategy Based on Intersection Signal Control*. Xi'an: Xidian University.
12. Xu, Yanfeng, Haiyan Yu, and Bing Su. 2012. Traffic flow distribution models based on time and path preference and inducement strategy. *Systems Engineering-Theory & Practice* 32(10): 2306–2314.
13. He, Zhengbin, and Wei Guan. 2013. Traffic state feedback guidance strategy under long update intervals. *Control and Decision* 28(7): 1046–1050.

14. Qin, H., H. Guan, and Y.J. Wu. 2013. Analysis of park-and-ride decision behavior based on decision field theory. *Transportation Research Part F: Traffic Psychology & Behaviour* 18 (5): 199–212.
15. Wang, Y., M. Papageorgiou, G. Sarros, and W.J. Knibbe. 2006. Real-time route guidance for large-scale express ring-roads. In *Intelligent Transportation Systems Conference*, 224–229. IEEE.
16. Li, Honglian. 2011. *Urban Multi-mode Traffic Assignment Based on Transfer Condition*. Beijing: Beijing Jiaotong University.

Study on Evaluation Method of Reconstruction and Extension Expressway Alignment Based on Operating Speed

Xin Ding, Jianyou Zhao, Xiaoyu Fu and Yang Zhao

Abstract With the rapid development of economy, early expressway in our country is facing with reconstruction and extension. The expressway alignment, as the skeleton of expressway, is an important guarantee for the safety of the vehicle driving. Therefore, road alignment evaluation is a major content in the extension of the expressway. In this paper, on the basis of evaluation method of expressway alignment, with the combination of domestic and foreign research experience, the expressway alignment is evaluated from three aspects: continuity, comfort and safety, and the alignment adaptability evaluation system is established through principal component analysis (PCA). At last, the applicability of the adaptability model is tested and verified for the alignment adaptability evaluation of Beijing Shijiazhuang expressway.

Keywords Expressway · Operating speed · Evaluation index · Alignment adaptability

X. Ding
Department of Transportation Expressway Management Office,
Luoyang 471000, Henan, China

J. Zhao (✉) · X. Fu
School of Automobile, Chang'an University, Xi'an 710064, Shanxi, China
e-mail: jyzhao@chd.edu.cn

X. Fu
e-mail: 1034867994@qq.com

Y. Zhao
School of Civil Engineering, Chang'an University, Xi'an 710064, Shanxi, China
e-mail: 36126091@qq.com

Introduction

Since 1990s, expressway construction in our country has a rapid development. By the end of 2014, the national expressway mileage has reached 111,900 km, ranking first in the world. Usually, the expressway operation period is about 20 years, and with the maturity of operational life, there will be a lot of expressway extension project. The core of expressway reconstruction and extension is to reevaluate the original alignment condition and to improve the original alignment index, which has a significant effect on the reconstruction and extension project. In this paper, the effect of operating speed to alignment evaluation is analyzed based on Beijing Shijiazhuang Reconstruction and Extension Expressway Project. In accordance with operating speed, the indexes and model of alignment evaluation are established to provide a theoretical guidance for the alignment evaluation to reconstruction and extension expressway.

Analysis on Alignment Evaluation Index

Basic Requirements of Expressway Alignment Design

Expressway alignment should be designed according to the grade and the function of highway by applying technical indicators correctly to ensure consistence and balanced of alignment, which means highway alignment can well adapt to efficiency, safety, comfort and continuous operation of vehicle.

Evaluation Method of Expressway Alignment

Evaluation methods of expressway alignment in China and abroad mainly include traffic safety evaluation, checking alignment design and the coordination of landscape along the route by expressway pivot chart and 3D model, evaluating expressway alignment comfort according the horizontal axis acceleration, evaluating expressway alignment by utilization rate and the relationship between the standard deviation of speed and accident rate. Operating speed is the main index of expressway alignment evaluation, which can reflect the actual operation situation of the vehicle on the road. Research on the operating speed in our country starts relatively late, but in recent years, research on alignment evaluation based on the operating speed rapidly rises. The major research evaluates the safety, continuity, harmony and comfort of expressway alignment based on the operating speed. A wealth of research theory with a high reference value has been accumulated in the evaluation of highway alignment.

Existing Research Review

Although much research has been done on the freeway alignment evaluation at home and abroad, there is a lack of research focusing on the adaptability of the comprehensive effect of expressway alignment and qualitative analysis method rather than analysis model based on the quantitative data is used in few studies. In addition, little research results have been reported at this point concerning alignment adaptability evaluation on reconstruction and extension expressway. There are many theoretical studies on the single index related to alignment adaptability, such as alignment continuity, safety, consistency, and so on, especially in the relation between operation speed and alignment. But the research on the adaptive alignment comprehensive evaluation system is little. In this paper, continuity, comfort and safety indexes of the expressway alignment are studied and the alignment adaptability evaluation system is established based on the operating speed.

Establishment of Alignment Adaptability Evaluation System

Selection of Evaluation Indicators

Alignment adaptability should be considered from three perspectives: continuity, comfort and safety. The operating speed V_{85} is the speed of the eighty-fifth percentage point on the characteristic points of the road when the traffic is in the free-flow state and the weather is good. In this paper, each single evaluation index closely related to the operating speed is selected as the basis of the alignment comprehensive evaluation system.

Continuity Evaluation Index

The alignment continuity is closely related to the operating speed of the road. $\Delta_{85}V$ is adopted in this paper to be continuity evaluation index. Firstly, speed difference of a single vehicle on adjacent sections is calculated, and then the eighty-fifth percentage of the cumulative speed is taken as $\Delta_{85}V$. The evaluation criterion of alignment continuity is shown in Table 1.

Table 1 Evaluation criterion of alignment continuity $|\Delta_{85}V|$

Operating speed difference (km/h)	Alignment continuity
$|\Delta_{85}V| < 10$	Good
$10 < |\Delta_{85}V| < 20$	Average
$|\Delta_{85}V| < 20$	Poor

For the road with average operating speed continuity, when conditions permit, the adjacent sections of the technical indicators can be adjust to achieve a good continuity; for the road with poor operating speed continuity, the horizontal and vertical alignment technology index is needed to be readjust on adjacent sections. This method eliminates the disadvantages of $\Delta_{85}V$ calculation method to a great degree.

Comfort Evaluation Index

The comfort of expressway alignment commonly evaluates by the acceleration value acting on the human body. The main factors include vertical acceleration, lateral acceleration and axial acceleration, among which horizontal and axial acceleration have the greatest impact on comfort.

The vertical acceleration of the vehicle is caused by the vertical curve under driving situation, but related research shows that the vertical acceleration has a small influence on the ride comfort. Therefore, the key evaluation indexes of comfort evaluation in this paper are horizontal and axial acceleration. The evaluation criterion is shown in Tables 2 and 3.

The axial acceleration value can reflect degree vehicle acceleration and deceleration under driving situation, and ride comfort reduces with the increase of acceleration value. Consulting the domestic and foreign research results, the evaluation criterion is shown in Table 3.

Safety Evaluation Index

There are many evaluation indexes of expressway alignment safety, from the view of the operating speed, the commonly used indexes include the speed variation coefficient GV, the relative difference of unit mileage speed VR. The evaluation criterion is shown in Tables 4 and 5.

Table 2 Comfort evaluation criterion based on horizontal acceleration

Horizontal acceleration value (m/s^2)	Alignment comfort
$a_h \leq 2.0$	Good
$2.0 < a_h < 3.0$	Average
$a_h \geq 3.0$	Poor

Table 3 Comfort evaluation criterion based on axial acceleration

Vehicle acceleration		Vehicle deceleration	
Axial acceleration value (m/s^2)	Alignment comfort	Axial acceleration value (m/s^2)	Alignment comfort
$a_s < 0.9$	Good	$|a_s| < 1.3$	Good
$0.9 \leq a_s < 1.2$	Average	$1.3 \leq |a_s| < 2.5$	Average
$a_s \geq 1.2$	Poor	$|a_s| \geq 2.5$	Poor

Table 4 Evaluation criterion of operating speed variation coefficient

Speed variation coefficient value (%)	Alignment safety
$G_v < 10$	Good
$10 \leq G_v < 20$	Average
$G_v \geq 20$	Poor

Table 5 Evaluation criterion of relative difference value of unit mileage speed (*VR*)

Relative difference value of unit mileage speed (1/km)	Alignment safety
$|VR| < 0.25$	Good
$0.25 \leq |VR| < 1.0$	Average
$|VR| \geq 1.0$	Poor

1. Speed variation coefficient value
2. Relative difference value of unit mileage speed

The difference of the operating speed on the adjacent sections reflects the safety of the alignment to some degree; the greater the difference of the speed variation is, the larger the potential safety hazard is.

Establishment of Alignment Adaptability Evaluation Model

There are five indexes to participate in the adaptability evaluation of the alignment. As the indexes emphasis on different aspect, alignment adaptability can be evaluated from different angles. In this paper, the alignment adaptability is considered as comprehensive form of the driving safety and comfort, continuity. Therefore, this paper introduces a comprehensive evaluation index—alignment adaptability coefficient A—to express the comprehensive evaluation results of individual indexes. The adaptation coefficient and the level of individual indexes are shown in Fig. 1.

In the comprehensive evaluation of each index, because the numerical unit of each index is variety, it is necessary to apply dimensionless method to the numerical value of different indexes. Since each index can be divided into three levels: good, average and poor. Fuzzy comprehensive evaluation method is used to assign each grade of evaluation index that each evaluation index is endowed with values between 0 and 1. Therefore, the effect caused by different units of data can be eliminated. The specific assignment is shown in Table 6.

For the sake of the single evaluation index, the β value of each section is calculated and the corresponding numerical value is obtained by the formula conversion.

Because the assignment is an interval, selecting with a poor accuracy, the evaluation results may be unsatisfactory. For each evaluation index, after the β
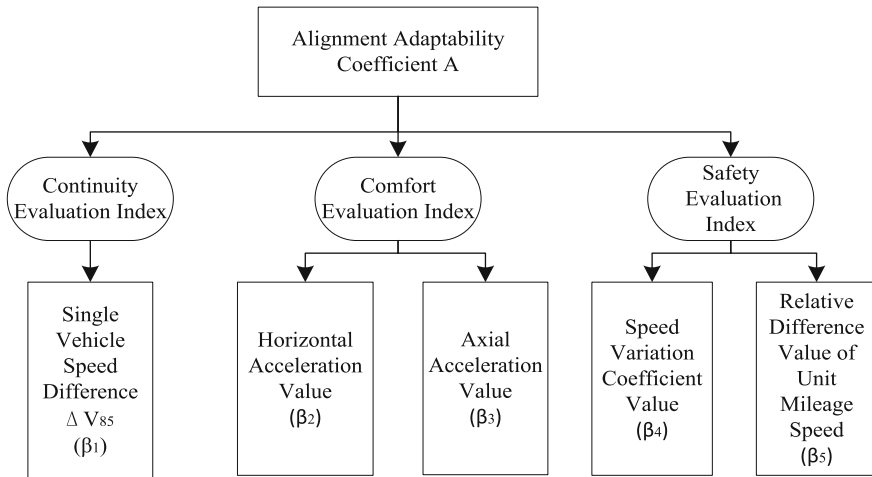
Fig. 1 Alignment adaptive evaluation index system

Table 6 Alignment adaptive evaluation index system

Evaluation level	Representative value	Recommended value
Good	0.8–1.0	0.9
Average	0.5–0.8	0.6
Poor	0.1–0.5	0.3

value of each divided sections has been computed by, the corresponding numerical can be obtained through Eq. 2.1. The equation is shown as follows:

$$\beta^* = a + \frac{\beta_{\max} - \beta}{\beta_{\max} - \beta_{\min}}(b - a) \tag{2.1}$$

where

β^* representative value after conversion of index;

β calculating value before conversion;

$\beta_{\max}$ upper bound of the index in a certain standard. When there is no upper limit value, the maximum value is the upper limit;

$\beta_{\min}$ lower bound of the index in a certain standard;

a, b critical value of assignment interval.

The above methods can improve the accuracy of assignment and are able to get more reasonable evaluation results.

The alignment adaptive equation is as follows:

$$A = \lambda_1 \beta_1 + \lambda_2 \beta_2 + \lambda_3 \beta_3 + \lambda_4 \beta_4 + \lambda_5 \beta_5 \tag{2.2}$$

where λ_1, λ_2, λ_3, λ_4 and λ_5 are the weighted values of evaluation index.

Single vehicle speed difference $\Delta V85$ $\beta 1$, horizontal acceleration value $\beta 2$, axial acceleration value $\beta 3$, speed variation coefficient value $\beta 4$, and relative difference value of unit mileage speed $\beta 5$ are the representative values of evaluation index.

Solution of Model Coefficients

The coefficient values of indexes in adaptability coefficient model that is the weight of each index. Usually, weight value is given by using expert scoring method or experience method directly, which can reflect a certain real situation, but due to the existence of the subjectivity, the evaluation results may not match the actual situation.

As each index is calculated based on the operating speed, there is a certain correlation between the indicators. Principal component analysis (PCA) is used to compute; on the basis of the correlation of the indexes, more original indexes are replaced by less comprehensive indexes through reducing the dimension; the characteristic value and the characteristic vector of the comprehensive index are related. When characteristic value reaches principal component contribution rate, the factor of feature vector can represent contribution degree of each index to comprehensive index and is the weight value of each index.

According to the existing survey data, the number of indexes has been determined; SPSS software is adopted to calculate and analyze using PCA method. The specific methods are as follows:

(1) Data calculation and standardization

In accordance with the velocity data of the survey, the corresponding value of each evaluation index is calculated, respectively, and the data matrix of five columns of N line is established by combining the number of evaluation indexes:

$$X = \begin{pmatrix} x_{11} & \cdots & x_{15} \\ \vdots & \ddots & \vdots \\ x_{m1} & \cdots & x_{m5} \end{pmatrix} \tag{2.3}$$

where m is the number of road sections;

X_{ij} data results of i_{th} section on j_{th} index.

The data unit of each index is standardized in use of processing method of the standard Z-score transform. The equation is shown as follows:

$$X'_{ij} = \frac{X_{ij} - \bar{X}_j}{S_j}(i = 1, 2 \ldots n; \quad j = 1, 2 \ldots 5) \tag{2.4}$$

where X'_{ij} is rata after standardized; X_{ij} raw data; $\bar{X}_j$ average value of j; and S_j standard deviation of j.

(2) Solution of the correlation coefficient matrix R of index data

The equation of the correlation coefficient matrix is as follows:

$$R = (r_{ij})_{n \times 5}; \quad r_{ij} = \frac{\text{cov}(x_i, x_j)}{\sqrt{DX_iDX_j}} \tag{2.5}$$

(3) Solution to the matrix R

By the relationship between characteristic value and characteristic vector, $Ra_i = \lambda_i a_i$ can be obtained, where λ_i and a_i are the corresponding characteristic value and unit characteristic vector. And $\lambda_1 \geq \lambda_2 \geq \cdots \geq \lambda_5 \geq 0$ should be satisfied.

(4) Solution to contribution rate of the principal component

The characteristic value is usually used to indicate the size of the principal component contribution rate. The equation is shown in Eq. 2.6:

$$G(k) = \sum_{i=1}^{k} \lambda_i / \sum_{i=1}^{5} \lambda_i (k \leq 5; \quad k \text{ is a round number}) \tag{2.6}$$

(5) Result analysis

Each factor of characteristic vector corresponding to the first principal component is considered as the weight value. In order to meet the adaptive coefficient A $(0 < A \leq 1)$, unit processing should be done. The factor of the unit vector after processing is the weight value of each index. Specially, if the first principal component contribution rate is not big, the second or the third principal component is needed. At this time, characteristic value and characteristic vector should be processed in weighted average method to achieve the final weight value.

Engineering Application

Jing-gang-ao Expressway grew out of Jingzhu national main trunk line, which is one of the earliest planning of the expressway in China. In recent years, with the rapid development of economy, traffic volume on the expressway increased

significantly, proportion of large vehicles is increasing continuously and road diseases are more and more serious. At the same time, its service levels also declined year by year. Therefore, the four-lane expressway has been unable to meet the travel needs of people. Beijing Shijiazhuang Expressway is a reconstruction and extension expressway. In this paper, the selected evaluation section is from K47+900 to K145+950. The alignment evaluation results of the road section are shown in Table 9

Data Acquisition

Because the whole line of Beijing Shijiazhuang section is long, it is not adaptive to collect road speed intensively. Thus, special section of Beijing Shijiazhuang Expressway in Jing-gang-ao Expressway is chosen to measure the operating speed. The measuring section mainly contains service area export, ramp exit and accident prone sections. Observation data of operating speed is measured by the fixed point video, radar gun and other methods. When collecting the speed, sample size is ensured more than 100 vehicles.

The calculation results of each index can be obtained by the operating speed of the vehicle, which is seen in Table 7. Standard index data is shown in Table 8. The calculation results of adaptability coefficient are shown in Table 9 horizontal acceleration value $\beta2$, axial acceleration value $\beta3$, speed variation coefficient value $\beta4$, relative difference value of unit mileage speed.

Calculation of Alignment Adaptive Coefficient

According to the calculated alignment adaptive coefficient when alignment adaptation coefficient A ranges from 0.8 to 1.0, road alignment adaptability is good; when adaptive coefficient A ranges from 0.5 to 0.8, road alignment adaptability is average; when adaptive numerical A ranges from 0.1 to 0.5, road alignment adaptability is poor. The results of adaptability evaluation of all sections are shown in Fig. 2.

With reference to Fig. 2, there are four sections with poor alignment adaptability. Combined with the actual survey, it is found that the sections with poor adaptability always are sections with entrances and exits of interchange or transition sections between long line and interchange. The alignment combination is not complicated, but for the sake of the wide vehicle speed change caused by excessive speed difference between adjacent sections, poor continuity of alignment, driving comfort and safety eventually lead to a poor alignment adaptability which reflects the comprehensive effect. The first seven and the 11th, 13th, 14th and 16th sections have good alignment adaptability, as these sections mainly have straight and large

Table 7 Calculation results of each evaluation index

Ordinal	Eighty-fifth percentage of speed difference $|\Delta_{85}V|$ (km/h)		Horizontal acceleration value $|a_h|$ (m/s²)		Axial acceleration value a_s (m/s²)		Speed variation coefficient value GV (%)		Relative difference value of unit mileage speed VR (1/km)	
	Passenger car	Truck	Passenger car	Truck	Passenger car	Truck	Passenger car	Truck	Passenger car	Truck
1	10.344	9.187	0.843	0.586	−0.248	−0.075	4.741	9.359	0.073	0.046
2	8.185	9.655	0.000	0.000	0.367	−0.141	3.541	7.460	0.107	0.094
3	10.795	10.985	2.083	1.492	0.251	0.130	2.475	5.389	0.073	0.092
4	11.278	8.681	0.000	0.000	0.517	−0.107	1.843	6.884	0.135	0.071
5	11.779	10.295	3.024	1.769	0.301	0.095	9.139	6.883	0.072	0.065
6	10.650	10.119	1.302	0.576	−0.336	−0.120	10.624	6.672	0.071	0.078
7	10.080	10.561	3.118	1.729	0.297	0.140	10.435	6.002	0.066	0.097
8	25.213	14.476	6.751	4.342	−1.408	−0.325	14.144	14.287	0.922	0.222
9	15.558	17.882	2.128	1.075	−0.874	−0.567	5.640	4.423	0.210	0.428
10	27.038	17.425	5.561	2.967	−2.967	0.975	15.972	21.323	0.901	0.693
11	7.080	7.489	0.000	0.000	−0.448	−0.010	7.145	18.113	0.138	0.011
12	20.840	14.038	4.335	2.304	0.986	0.902	18.599	20.638	0.969	0.373
13	18.176	21.127	4.231	2.116	0.874	0.823	17.431	19.424	0.037	0.087
14	5.090	23.240	0.000	0.000	−0.336	−0.021	7.120	15.324	0.254	0.124
15	9.763	15.378	0.824	0.546	−0.244	−0.066	4.621	8.943	0.196	0.223
16	19.436	17.188	6.252	4.124	−1.318	−0.334	13.334	13.442	0.058	0.064
17	8.774	12.748	3.021	1.744	0.322	0.087	8.933	6.782	0.207	0.245

Table 8 Data after standardization of each evaluation index

Ordinal	Eighty-fifth percentage of speed difference $|\Delta_{85}V|$(km/h)		Horizontal acceleration value $|a_h|$(m/s²)		Axial acceleration value a_s (m/s²)		Speed variation coefficient value GV (%)		Relative difference value of unit mileage speed VR (1/km)	
	Passenger car	Truck	Passenger car	Truck	Passenger car	Truck	Passenger car	Truck	passenger car	Truck
1	0.790	0.816	0.906	0.935	0.962	0.989	0.905	0.813	0.942	0.963
2	0.836	0.827	1.000	1.000	0.918	0.978	0.929	0.851	0.914	0.925
3	0.776	0.770	0.774	0.834	0.944	0.971	0.951	0.892	0.942	0.926
4	0.762	0.826	1.000	1.000	0.885	0.983	0.963	0.862	0.892	0.943
5	0.747	0.791	0.685	0.803	0.933	0.979	0.817	0.862	0.942	0.948
6	0.781	0.796	0.855	0.936	0.948	0.982	0.781	0.867	0.943	0.938
7	0.798	0.783	0.676	0.808	0.934	0.969	0.787	0.880	0.947	0.922
8	0.204	0.666	0.100	0.562	0.773	0.950	0.676	0.621	0.531	0.822
9	0.632	0.564	0.769	0.881	0.866	0.913	0.887	0.913	0.832	0.729
10	0.100	0.577	0.372	0.691	0.383	0.506	0.621	0.100	0.540	0.623
11	0.858	0.850	1.000	1.000	0.931	0.998	0.857	0.557	0.890	0.991
12	0.251	0.679	0.562	0.753	0.714	0.554	0.542	0.114	0.502	0.751
13	0.321	0.624	0.544	0.676	0.692	0.621	0.623	0.122	0.495	0.812
14	0.844	0.820	1.000	1.000	0.897	0.924	0.748	0.566	0.871	0.941
15	0.122	0.556	0.334	0.672	0.291	0.546	0.622	0.100	0.552	0.633
16	0.212	0.673	0.128	0.424	0.745	0.833	0.664	0.531	0.544	0.791
17	0.699	0.774	0.688	0.827	0.956	0.781	0.722	0.861	0.970	0.933

Table 9 Alignment adaptive coefficient of sections

Ordinal	Zero pile	Terminal pile	Alignment adaptive coefficient	
			Passenger car	Truck
1	K47+900	K52+825	0.893	0.889
2	K52+825	K58+170	0.902	0.894
3	K58+170	K63+770	0.882	0.874
4	K63+770	K69+390	0.874	0.901
5	K69+390	K70+370	0.842	0.879
6	K70+370	K75+480	0.863	0.887
7	K75+480	K85+070	0.852	0.872
8	K85+070	K91+000	0.477	0.740
9	K91+000	K110+350	0.785	0.761

Ordinal	Zero pile	Terminal pile	Alignment adaptive coefficient	
			Passenger car	Truck
10	K110+350	K114+010	0.380	0.481
11	K114+010	K118+460	0.893	0.858
12	K118+460	K128+940	0.485	0.558
13	K128+940	K133+550	0.848	0.757
14	K133+550	K138+660	0.843	0.771
15	K138+660	K142+150	0.437	0.564
16	K142+150	K145+950	0.831	0.782
17	K145+950	K150+700	0.762	0.725
18	K150+700	K153+000	0.781	0.733
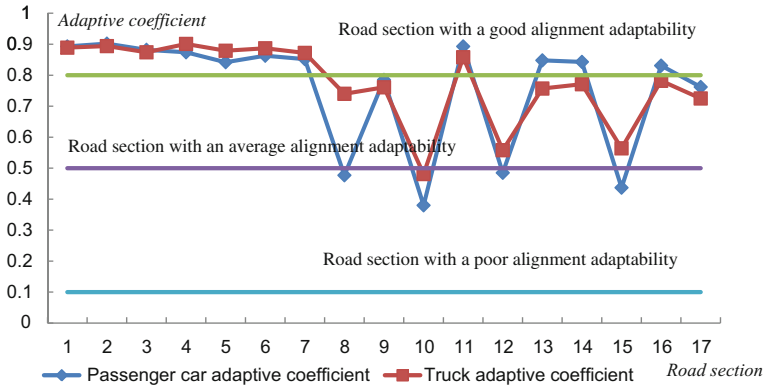
Fig. 2 Evaluation results of road alignment adaptability

radius curve and the maximum longitudinal slope is 1.73%, with the longitudinal slope of whole sections below 3%. The simple, coordinated alignment condition gives the vehicle a very smooth ride and then the comfort and safety of driving is ensured, therefore, the sections have the good alignment adaptability.

Conclusions

In this paper, an empirical analysis is done on the basis of the data of Beijing Shijiazhuang Reconstruction and Extension Expressway Project. In accordance with the operating speed variation model, the continuity index value of evaluation section is calculated; the expressway alignment is evaluated from three aspects: continuity, comfort and safety; and the value of safety and comfort evaluation index and weight value of indexes are obtained according to evaluation criteria. Finally, an adaptability evaluation to a road section of Beijing Shijiazhuang Reconstruction and Extension Expressway is done through the established evaluation model to test and verify the alignment adaptability evaluation system.

References

1. Zhang, Wenzhang. 2000. *Practical modern statistical analysis method and SPSS application*. Beijing: Contemporary China Publishing House.
2. Jin, Yakun. 2015. Study on the evaluation method of road alignment comfort. *Traffic world (Engineering Technology)* 07: 140–141.
3. Sheng, Lin. 2014. *Highway safety design theory research based on linear design consistency*. Beijing: Beijing Jiaotong University.

4. Hongzhi, Yang, Jingtao, Zhang, Jinliang, Xu. 2010. Two lane highway plane alignment safety evaluation based on the running speed. *Highway Traffic Science and Technology* 09:127–131 +142.
5. Xiaofei, Sui. 2012. *Evaluation and improvement measures of high grade highway alignment based on operation speed*. Xi'an: Chang'an University.
6. Tang, Qiusheng, and Xuguang Chen. 2014. Evaluation of highway alignment coordination based on operation speed. *Journal of Shandong University of Technology (Natual Science Edition)* 06: 39–43.
7. Cui, Ranran. 2013. *Research on highway design consistency evaluation based on the running speed*. Harbin: Harbin Institute of Technology.

Transfer Passenger Distribution Prediction on Flow Lines in Transportation Terminal

Siyuan Shao, Chunfu Shao and Wei Jin

Abstract Transportation terminal, as an important place that connecting to inner city and outer city traffic, can realize passenger transfer. The main parameters of the transportation terminal, transfer passenger flow, and its distribution also have important significance to the operational organization. According to the survey data in Beijing South Railway Station, this paper obtains transfer volume between transportation modes by calculating and calibrating impedance function. Then a flow line model based on SUE is established, and the flow lines are precisely measured by completing the distribution of transfer passenger. Finally, the paper provides a reasonable transfer organization scheme.

Keywords Transfer passenger · Flow distribution · Beijing South Railway Station

Introduction

Transportation terminal, as an important place that connecting to inner city and outer city traffic, can realize passenger transfer through the cohesion of a variety of transportation modes and giving full play to terminal facilities functions. With the growth of passenger volume and the increase of transfer demand of transportation modes, the objective and accurate analysis on passenger volume and transfer passenger flow of the terminal has become the important reference of improving the efficiency of passenger travel, standardizing the terminal operation management, and designing a new terminal.

Overseas researches on passenger travel and transfer behaviors of the transportation terminal are early. Disaggregate model is proposed based on random utility theory, which is applied to transportation division problem [1]. Chinese scholars have also done a lot of researches on the passenger volume forecasting of

S. Shao · C. Shao (✉) · W. Jin
School of Traffic and Transportation, Beijing Jiaotong University,
Beijing, China
e-mail: cfshao@bjtu.edu.cn

the transportation terminal [2–4]. However, the problem of lack of basic information, poor systematic, more qualitative research, and less operational qualitative analysis is still existent.

This paper takes Beijing South Railway Station as an example. It analyzes the complexity of transfer passenger flow within the transportation terminal and establishes a passenger flow line model. Through predicting the transfer volume in each transportation mode of the terminal and the distribution of transfer passenger flow on each flow line in the station, the flow lines are precisely measured and the bottleneck of the system can be found. Meanwhile, optimization measures are proposed.

Forecasting Method

The technical route is shown in Fig. 1.

Traffic Assignment Model

Transfer flow lines in the terminal are like a road network in the city, and the transfer volume between transportation modes can be seen as OD traffic volume in

Fig. 1 Technical route
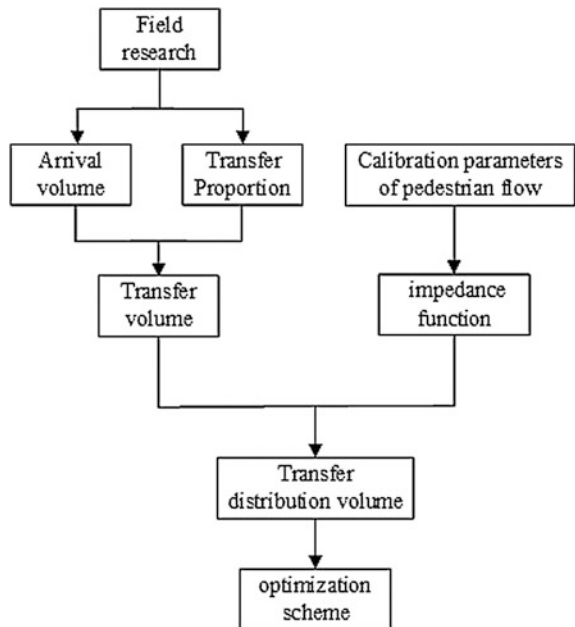
each district. Using traffic assignment theory of urban traffic, transfer distribution volume on each flow line can be calculated by distributing transfer passenger flow.

Equilibrium assignment approach takes effect of travel time on link flow into consideration. Thus link flow and travel time keep consistent in calculation. Though iterations lead to large compute, the result is still superior to most assignment methods. Stochastic user equilibrium (SUE) model assumes that travelers fail to acknowledge the complete information of the path in the network, or have diffident feelings about the information of the network, and the result will be closer to the reality than UE model. This is because SUE model allows travelers not only to choose the attractive paths but also those less attractive paths. So zero flow link would not occur in this situation.

SUE model is as follows [5]:

$$\text{min: } Z(x) = -\sum_{r,s} q_{rs} S[c^{rs}(x)] + \sum_a x_a t_a(x_a) - \sum_a \int_0^{x_a} t_a(w)\, dw, \quad (1)$$

where $S[c^{rs}(x)]$ is the satisfaction function, defined as the expected perceived travel cost; x_a is the flow of link a; t_a is a link cost function for link a; and q_{rs} is the demand for OD pair (r, s).

Required Data for Forecasting

Necessary input data for the traffic assignment model include a road network with properties, an OD matrix and impedance function. In this paper, the road network is all of transfer flow lines. The OD matrix is the transfer volume between transportation modes, which can be calculated by multiplying the following two parts. One is arrival passenger volume of each transportation mode which can be got according to relevant information such as traffic operation plan, vehicle parameter data, supplementary survey data, and so on. The other is transfer proportion between each transportation mode, which can be achieved by field research or referring to the similar type of the transportation terminal. The classic BPR function is chosen as impedance function.

Case Study

Passenger Flow Analysis

Through an investigation in Beijing South Railway Station on April 22, 2014, the data of the arrival passenger volume from 15:30 to 17:30 were collected, as shown in Fig. 2. The figure shows that the arrival volume from the railway has a large
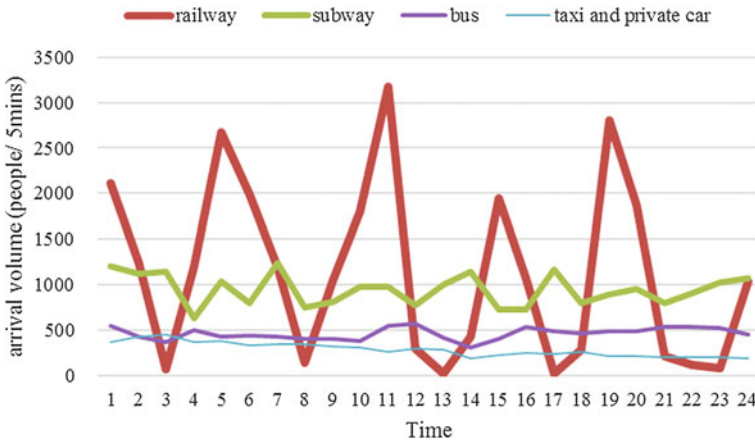
Fig. 2 Arrival passenger volume

proportion of the total arrival volume, and the value increases rapidly to form a passenger flow peak as a train arrives. Compared to the railway, the subway has a more uniform passenger flow due to its smaller departure interval. The proportion of the arrival volume from buses, taxis, and private cars is relatively smaller, and the passenger flow has little change.

In addition, by sampling questionnaire survey of passenger flow and referring to the similar type of the transportation terminal, the transfer proportion can be obtained. Then, taking one hour from 16:00 to 17:00 as the peak period of passenger volume in the terminal, the transfer volume can be calculated, as shown in Table 1.

Passenger Flow Assignment on Transfer Flow Lines

First, a comprehensive analysis on each layer in Beijing South Railway Station is done. Figure 3 shows the network model of passenger flow lines in the terminal.

Then the parameters of BPR function are calibrated by the measured data of pedestrian flow in Beijing South Railway Station, as shown in Table 2.

Referring to the capacity of each channel facility in Code for Design of Metro [6], as shown in Table 3 for details, link parameters of the flow lines are input, as shown in Table 4.

Figure 4 shows the outcome on passenger flow assignment using TransCAD software.

Table 1 The transfer volume between different transportation modes

Transportation mode	Beijing–Tianjin Railway	Beijing–Shanghai Railway	Other Railway	Subway	Bus	Taxi and private cars
Beijing–Tianjin Railway	0	231	231	2543	925	694
Beijing–Shanghai Railway	86	0	86	948	345	259
Other Railway	354	354	0	3896	1417	1063
Subway	3190	2658	1914	0	2658	213
Bus	1074	1074	537	2041	537	107
Taxi and private cars	1342	1342	632	326	326	0

Note Line 14 was not opened at the time of investigation, so this paper does not consider the influence of Line 14 on the terminal
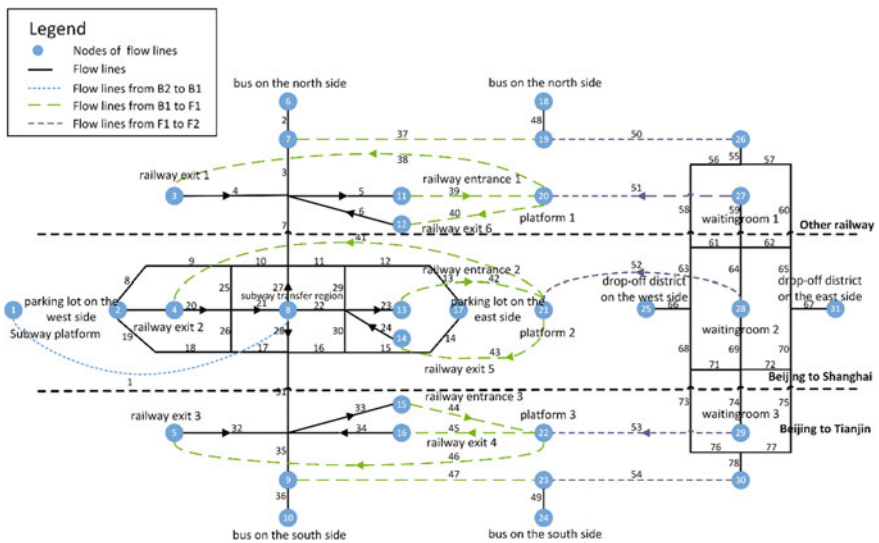
Fig. 3 Flow lines and link numbers

Table 2 Calibration results of the impedance function in the terminal

Passageway types	Parameter a	Parameter β
Passageway	1.42	5.0365
Escalator	0.45	1.4021

Table 3 The capacity of each channel facility

Types	Capacity (people/h)
Passageway with 1 m wide	4000
Escalator with 1 m wide	7200
One fare gate	2100

Table 4 Parameters of the flow lines

Number	Link types	Length (m)	Average width (m)	Speed (m/min)	Free flow travel time (min)	Capacity (people/h)
1	B2 escalator	30	1	30	1	14,400
2	B1 passageway	40	4	60	0.67	16,000
3	B1 passageway	60	15	60	1	60,000
...						
21	B1 passageway and gates	10	4	60	0.17	8400
22	B1 passageway and gates	10	4	60	0.17	10,500
...						
27	B1 passageway and gates	10	4	60	0.17	8400
28	B1 passageway and gates	10	4	60	0.17	8400
...						

Note The capacity of Link 1 is 14,000 people/h because it has two escalators. If the link is equipped with fare gates, the capacity will choose the lower on, such as Link 21
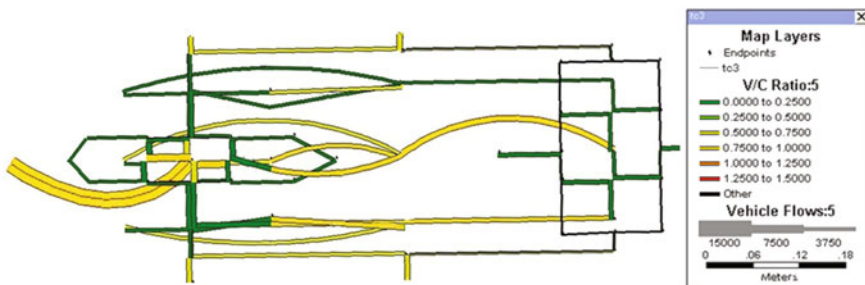
Fig. 4 Passenger flow assignment

Table 5 Results of flow distribution in the subway transfer region

Number	Link types	Traveling direction	East (south) line	West (north) line	Saturation (%)
21	B1 passageway and gates	East–west	5982	0	0.7121
22	B1 passageway and gates	East–west	2173	4486	0.6342
27	B1 passageway and gates	North–south	0	4460	0.5310
28	B1 passageway and gates	North–south	4257	0	0.5068

Table 6 Results of flow distribution improved in subway transfer region

Number	Link types	Traveling direction	East (south) line	West (north) line	Saturation (%)
21	B1 passageway and gates	East–west	4753	0	0.5658
22	B1 passageway and gates	East–west	0	5815	0.5538
27	B1 passageway and gates	North–south	0	5541	0.5277
28	B1 passageway and gates	North–south	5310	0	0.5057

Optimization on Flow Lines

By the assignment results, there are a large number of passengers in and out of the subway, which is the bottleneck area in the basement level for evacuation. Table 5 lists the results of flow distribution in the subway transfer region.

Due to the large passenger flow in the subway transfer legion, the saturation of Link 21 is high. While Link 22 is bidirectional, it is inevitable that flow lines of passenger entering and departing will be crossing. Therefore, a one-way entrance should be considered to deal with the problem. In order to ensure that passengers can be successfully out of the subway, trying to open one more fare gates in the north and south side under present conditions to increase the capacity of links, the distribution results are better, as shown in Table 6.

Conclusion

Predicting the transfer distribution of transportation modes in a scientific way, which has guiding significance to the construction planning and operational organization of the transportation terminal, this paper establishes a model of passenger

flow lines to predict transfer distribution based on SUE model and calibrated impedance function, and provides a basis to determine a reasonable transfer organization scheme by completing the distribution of transfer passenger in Beijing South Railway Station.

Due to the limitation of the data, the transfer volume forecast does not consider passengers' personal attributes, travel preferences, and other factors. It is failed to use multinomial logit model to calculate the transfer proportion. Further research should be continued to improve the accuracy of the model.

References

1. Denhardt, Robert B., Janet Vinzant Denhardt. 2000. The new public service: Serving rather than steering. *Public Administration Review* 60(6): 549–559.
2. Yingnan, Wang. 2008. *Research on Passenger Flow Forecast and Optimal Method of Organization in Comprehensive Transportation*. Beijing: Beijing Jiaotong University.
3. Wang, Sujuan. 2012. *The Transfer Volume Forecasting on Integrated Passenger Transportation Hub*. Jilin: Jilin University.
4. Zhao, Lin, Duan Tiezheng, Haitao Yu, et al. 2014. Analysis of Prediction model for Urban Public Transportation Hub. In *Conference on Urban Development and Planning, Tianjin*, 23–24. September 2014.
5. Chunfu, Shao. 2014. *Traffic Planning*. Beijing: China Railway Publishing House.
6. GB 50157-2013. Code for Design of Metro.

Critical Safety Distance Model of Human-Vehicle Unit Based on Traffic Conflict at Urban Intersections

Fang Li, Le-yi Wang, Wu-hong Wang and Xiao-bei Jiang

Abstract Traffic safety problem is one of the main problems faced by road traffic system, which attracts more and more attention of the society. Intersection is the bottleneck of the urban road traffic system. A large proportion of traffic accidents take place at intersections. On the basis of the existing critical conflict zone model, the human-vehicle unit driving states were further divided into longitudinal driving state, lateral driving state, and intersectional driving state, and traffic conflicts at intersections were analyzed. Emphasis was put on traffic conflict danger potential reaction of the human-vehicle unit in intersectional driving state. Combined with traffic rules, the risk identification model based on traffic conflict in intersection condition was built. Calculation examples show that the model can well simulate the actual traffic conditions and improve the precision and practicability of the original model. The theory of driver behavior and traffic safety has also been expanded and enriched.

Keywords Traffic safety evaluation · Traffic conflict · Critical conflict zone · Driver behavior

Introduction

With the rapid development of urbanization and motorization, traffic safety, as a more and more important problem, is attracting universal attentions of the society. Exploration on how to reduce the number of accidents and improve the traffic safety becomes the most important research subject in traffic area.

Prevention is the most effective risk control measure. Risk Identification Theory (RIT) is able to improve traffic safety from the active prevention level, which also becomes the research hot spot. Traffic safety evaluation is an effective method to identify traffic accident risk and improve traffic security. At present, most of the

F. Li · L. Wang · W. Wang (✉) · X. Jiang
School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China
e-mail: wangwuhong@bit.edu.cn

traffic safety evaluation models and methods are based on statistical analysis of the incident data, which cannot reveal the accident forming mechanism in detail. However, the traffic conflict evaluation method does not depend on traffic accident data, and its theory is closer to the intrinsic reason of the accidents, which has been widely applied.

Critical conflict zone model based on vehicle kinematics and traffic conflict technique is a method to discriminate traffic conflicts. Guo [1] gave the definition of critical conflict zone model. The study shows that the risk perception scopes of the traffic participants, such as motor vehicles, non-automatic vehicles, and pedestrians, are not confined to a specific direction or a certain distance, but to a certain region within the scope of the risk perception ability. The critical conflict zone was quantified by variables such as velocity, distance, and angle. Therefore, the discriminant model of critical conflict zone based on vehicle dynamics and traffic conflict technique shows more practical application value [1]. Wang [2] improved the critical conflict zone model, in consideration of the effect of the size of the vehicles on the conflicts, then analyzed the parameters in the model, and observed the effect of the parameters on the conflict radius. According to the previous research results, the critical conflict zone model has the advantages of high precision and simple operation. But it also has some deficiencies: Firstly, it leaves driver's reaction to the potential conflict danger during the driving process out of consideration; secondly, it ignores that the traffic rules affect the actual vehicle trajectories. In this paper, these two points will be corrected to improve the critical conflict zone model.

In order to quantify the critical conflict zone, some assumptions are set as follows:

(1) The sizes of vehicle A and B on urban road are ignored;
(2) Left- and right-turning actions at intersections are simplified to linear motions.

Analysis of Driver Behavior

Urban road traffic can be divided into road traffic and intersection traffic. And longitudinal motion and lateral motion are the two typical vehicle movement forms in road traffic.

In longitudinal motion, a certain distance must be kept between the leading vehicle and the following vehicle. When there is a sudden brake made by the leading vehicle, if the following vehicle cannot take some safety measures, such as braking, to avoid the emergency timely, there will be a collision.

In lateral motion, a certain distance must be kept between the target vehicle and the vehicle behind it in the target lane, as there is a conflict between them. When the conflict happens, if the behind one in the target lane cannot do some safety measures, such as braking, to avoid the emergency timely, there will be a collision.
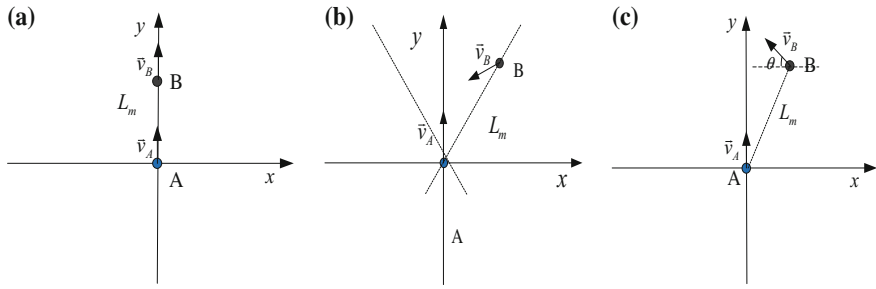
Fig. 1 Driver behavior abstract sketch: **a** longitudinal motion, **b** intersectional motion, and **c** lateral motion

At intersections, there are serious interferences between motor vehicles and pedestrians crossing the street, motor vehicles and motor vehicles, and non-automatic vehicles and motor vehicles, which lead to greatly increased traffic flow compared to road traffic. Otherwise, as the traffic flows drive to various directions, there are lots of conflicts between traffic participants, easy to cause traffic accidents.

Vehicle motions on urban roads can be abstracted as the following three conditions:

By integrating (a), (b), and (c) in Fig. 1, we can get the critical conflict zone model of the human-vehicle unit, as shown in Fig. 2.

Considering that the safety distances in longitudinal motion and lateral motion have already been confirmed [3], only the critical conflict zone in intersectional motion will be discussed in this paper.

Critical Conflict Zone Model at Intersections

Conflict Analysis at Intersections

Intersections usually have three forms: T junction, Y junction, and crossing, while T junction and Y junction can be seen as special forms of crossing. Hence, only crossing will be analyzed in the paper, which includes diverging points, merging points, and conflict points, as shown in Fig. 3.

Diverging points, merging points, and conflict points all have potential conflicts. Diverging conflicts involve longitudinal motion in the same lane and lateral motion in two adjacent lanes, which are similar to the traffic states in longitudinal motion and lateral motion, and the diverging state has less danger, so in this paper, only conflicts in merging points and conflict points will be discussed.

Fig. 2 Critical conflict zone model
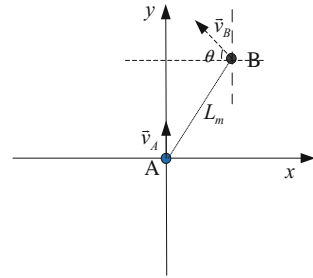
Fig. 3 Potential collision points at the intersection
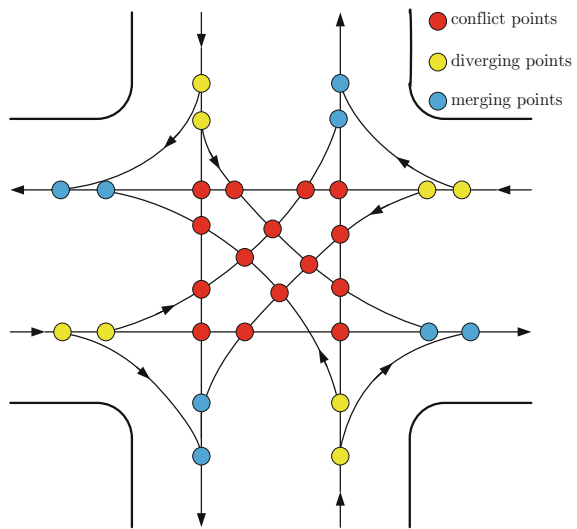
Critical Conflict Distance Model in Conflict Points

As shown in Fig. 4, there are 16 conflict points in the intersection. Among the 4 conflict points in each straight-going direction, the conflict points formed by the left-turning traffic flow from the adjacent direction and the opposite direction are the most important ones, which will be emphatically analyzed in this paper.

(1) Critical conflict distance model of straight-going vehicles and left-turning vehicles from opposite directions:

In intersection traffic, straight-going vehicle A encounters vehicle B which intends to turn left. According to the law of the People's Republic of China on road traffic safety, in non-signal intersections, vehicles turning left should give way to the vehicles going straight. Thus, in conditions shown in Fig. 5, vehicle B should give way to vehicle A. As a result, when vehicle B discovers vehicle A going

Fig. 4 Conflict points in the crossing
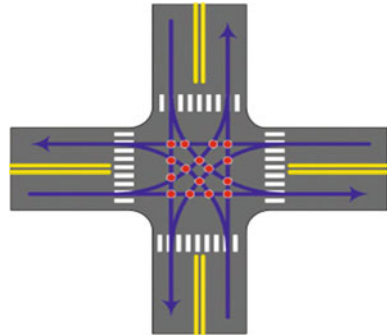
Fig. 5 Conflict between straight-going vehicle and left-turning vehicle from opposite directions
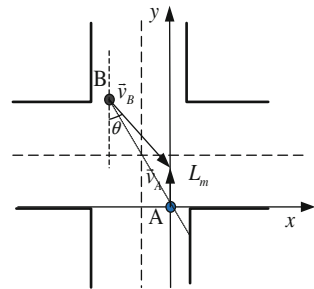
Fig. 6 Critical conflict between straight-going vehicle and left-turning vehicle from opposite directions
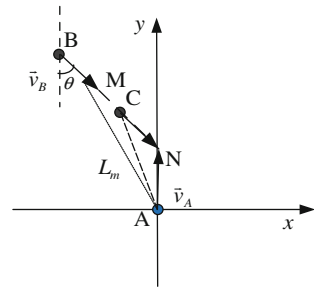
straight, vehicle B should brake to avoid vehicle A, till vehicle A has passed. Then, vehicle B continues to turn left.

As shown in Fig. 6, vehicle B gives way to vehicle A. Vehicle A moves at a constant speed and drives along line AN. The critical state is when vehicle B drives to point N and vehicle A has just passed the same point.

Driving distance of vehicle A is as follows:

$$x_A = v_A t, \quad t = \frac{v_B}{a} + t_0 \tag{1}$$

Fig. 7 Conflict between straight-going vehicle and left-turning vehicle from adjacent directions

Driving distance of vehicle B is as follows:

$$x_B = v_B t_0 + \frac{v_B^2}{2a} \tag{2}$$

In which, t_0 means the total time from the driver is aware of the danger to the brake pedal begins to work.

According to the triangle formula, the critical conflict distance of straight-going vehicle A and left-turning vehicle B from opposite directions can be described as follows:

$$L_m = \sqrt{x_A^2 + x_B^2 + 2x_A x_B \cos(\pi - \theta)} \tag{3}$$

(2) Critical conflict distance model of straight-going vehicles and left-turning vehicles from adjacent directions:

Conflicts between straight-going vehicles and left-turning vehicles from adjacent directions are restricted with the same regulations as the conflicts between straight-going vehicles and left-turning vehicles from opposite directions, and the only difference lies in the value of angle θ (Fig. 7).

The conflicts of straight-going vehicles and left-turning vehicles from adjacent directions also have the same analysis process with the conflicts of straight-going vehicles and left-turning vehicles from opposite directions. According to the triangle formula, the critical conflict distance of straight-going vehicle A and left-turning vehicle B from adjacent directions can be described as follows:

$$L_m = \sqrt{x_A^2 + x_B^2 + 2x_A x_B \cos\left(\frac{\pi}{2} + \theta\right)} \tag{4}$$

Critical Conflict Distance Model in Merging Points

As shown in Fig. 8, there are 8 merging points in a crossing. And each straight direction has two merging points, which are the merging point formed by the left-turning flow from the adjacent direction and the merging point formed by the right-turning flow from the adjacent direction.

(1) Critical conflict distance model of straight-going vehicles and right-turning vehicles from adjacent directions

In intersection traffic, when vehicle A in the straight direction encounters vehicle B which intends to turn right, according to the law of the People's Republic of China on road traffic safety, at non-signal intersections, vehicles going straight should give way to the vehicles turning right. Thus, in conditions shown in Fig. 9, vehicle A should give way to vehicle B. As a result, when vehicle A discovers vehicle B turning right, vehicle A should brake to avoid vehicle B, till vehicle B has passed, and then, vehicle A continues to go straight.

As shown in Fig. 10, vehicle A gives way to vehicle B. Vehicle B moves at a constant speed to drive along line BM. The critical state is when vehicle A drives in a constant speed to point N, vehicle B just arrives at the same point.

Considering that vehicle A and vehicle B will drive into the same lane after the merging point N, vehicle A would be better to keep the same speed with vehicle B and not to stop at point N. Driving distance of vehicle A is as follows:

Fig. 8 Merging points at the intersection
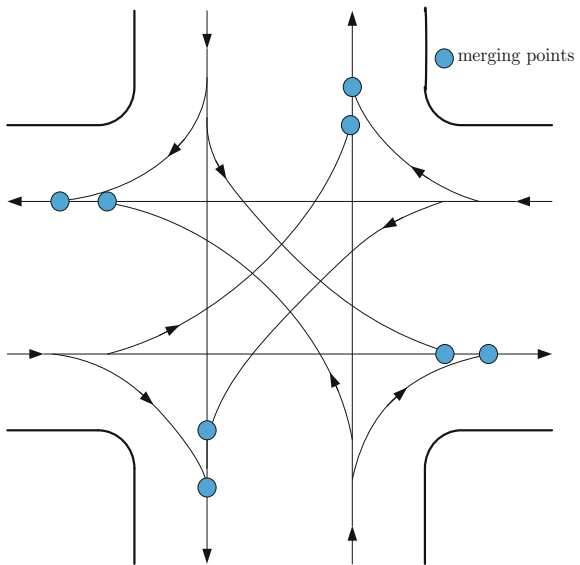
Fig. 9 Conflict between
straight-going vehicle and
right-turning vehicle from
adjacent directions
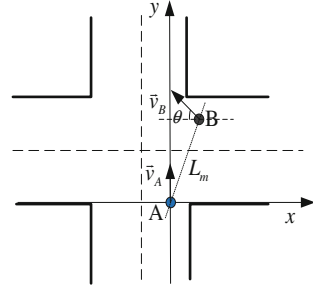
Fig. 10 Critical conflict
between straight-going
vehicle and right-turning
vehicle from opposite
directions
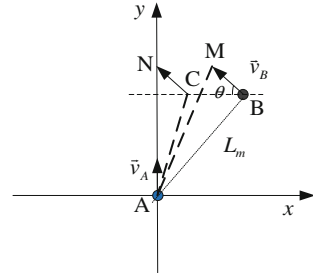
$$x_A = v_A t_0 + \frac{v_A^2 - v_B^2}{2a} \tag{5}$$

As vehicle B is a right-turning vehicle and drives at a low constant speed, driving distance of vehicle B is as follows:

$$x_B = v_B t, \quad t = \frac{v_A - v_B}{a} + t_0 \tag{6}$$

According to the triangle formula, the critical conflict distance of straight-going vehicle A and right-turning vehicle B from adjacent directions can be described as follows:

$$L_m = \sqrt{x_A^2 + x_B^2 + 2x_A x_B \cos\left(\frac{\pi}{2} - \theta\right)} \tag{7}$$

(2) Critical conflict distance model of straight-going vehicles and left-turning vehicles from adjacent directions

According to the law of the People's Republic of China on road traffic safety, at non-signal intersections, vehicles turning left should give way to vehicles going straight. Thus, in conditions shown in Fig. 11, vehicle B should give way to vehicle A. As a result, when vehicle B discovers vehicle A going straight, vehicle B

Fig. 11 Conflict between
straight-going vehicle and
left-turning vehicle from
adjacent directions
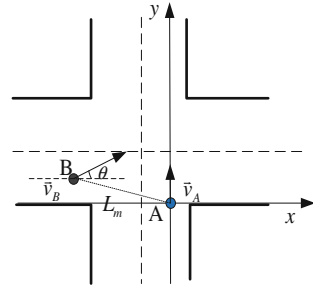
should brake to avoid vehicle A, till vehicle A has passed, and then, vehicle B continues to turn left and drive in the same line with vehicle A.

The analysis process is the same with the conflicts between straight-going vehicles and right-turning vehicles from adjacent directions. According to the triangle formula, the critical conflict distance of straight-going vehicle A and left-turning vehicle B from adjacent directions can be described as follows:

$$L_m = \sqrt{x_A^2 + x_B^2 + 2x_A x_B \cos\left(\frac{\pi}{2} - \theta\right)} \tag{8}$$

Calculation Examples

In this paper, small cars have been chosen as the research objects. According to Chinese vehicle design standard (GB7258-87) and motor vehicle inspection specifications made in 1978 by the highway research institute of the ministry of transport, the value of maximum braking deceleration α_{max} is 7.4 m s^{-2} and t_0 is 1.2 s. Usually, vehicle steering angle is within a certain range. In this paper, we make $\theta = \pi/6$.

(1) Calculation examples of the conflicts on conflict points

Based on the conflict analysis on conflict points, critical distances in the two critical distance models can be confirmed by the speeds of the vehicles and the identified parameters above. Fix the speed of vehicle B, and change the speed of vehicle A to obtain the results in Table 1 and Fig. 12.

As shown in Fig. 12, when the speed of vehicle B and other parameters are fixed, there tends to be a linear correlation between the critical conflict distance and the speed of vehicle A. However, in actual traffic state, parameters such as steering angle cannot be fixed. Therefore, the critical safe distance will be more complex.

Table 1 Critical distance on conflict points

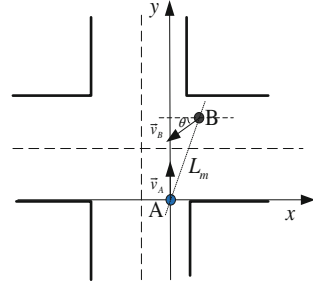| v_A (km/h)	v_B (km/h)	α (m/s^2)	Vehicle steering angle (rad)	Critical distance of straight-going vehicles and left-turning vehicles from opposite directions (m)	Critical distance of straight-going vehicles and left-turning vehicles from adjacent directions (m)
30	30	7.4	$\pi/6$	9.9	17.5
40	30	7.4	$\pi/6$	15.0	22.5
50	30	7.4	$\pi/6$	20.9	28.0
60	30	7.4	$\pi/6$	27.1	33.9
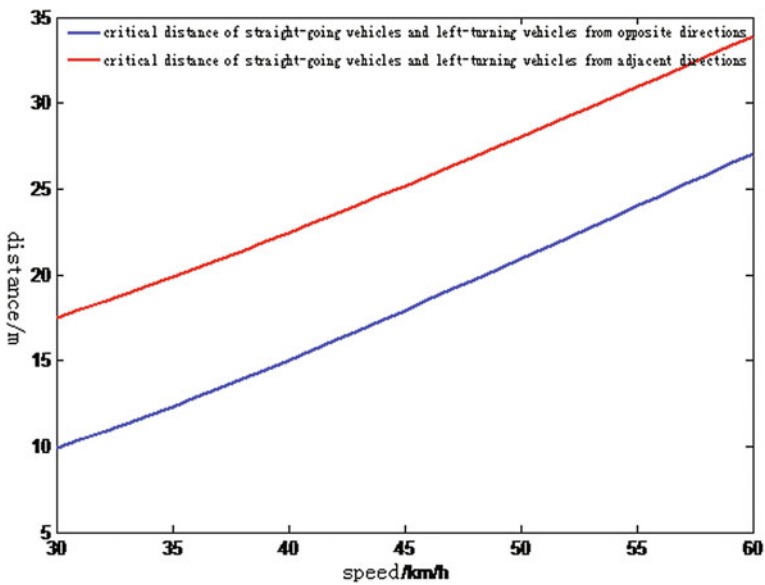
Fig. 12 Critical distance and speed of vehicle A on conflict points ($v_B = 30$ km/h)

Table 2 Critical distances on merging points

v_A (km/h)	v_B (km/h)	α (m/s^2)	Vehicle steering angle (rad)	Critical distance of straight-going vehicles and right-turning vehicles from adjacent directions (m)	Critical distance of straight-going vehicles and left-turning vehicles from adjacent directions (m)
30	30	7.4	$\pi/6$	17.3	17.3
40	30	7.4	$\pi/6$	26.1	27.0
50	30	7.4	$\pi/6$	36.0	39.6
60	30	7.4	$\pi/6$	46.9	54.9
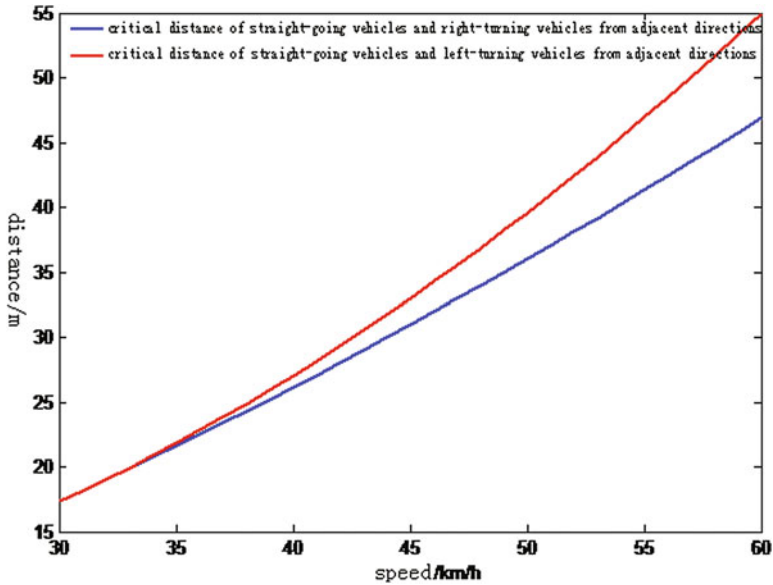
Fig. 13 Critical distance and speed of vehicle A on merging points ($v_B = 30$ km/h)

(2) Calculation examples of conflicts on merging points

Similar to the conflict analysis on conflict points, the critical conflict distances of the two critical conflict distance models on merging points can be confirmed with the speeds of the vehicles and the identified parameters above. Fix the speed of vehicle B, and change the speed of vehicle A to obtain the results in Table 2 and Fig. 13.

As shown in Fig. 13, when the speed of vehicle B and other parameters are fixed, there tends to be a linear correlation between the critical conflict distance and speed of vehicle A. However, in actual traffic state, parameters such as steering angle cannot be fixed. Therefore, the critical safe distance will be more complex.

Conclusions

On the basis of the existing critical conflict zone model, this paper further divides the driving states in urban road traffic into longitudinal driving state, lateral driving state, and intersectional driving state and analyzes the traffic conflicts at urban intersections. This paper puts emphasis on traffic conflict potential danger reaction of the human-vehicle unit in intersectional driving state. Combined with the traffic rules, the critical conflict distance model based on traffic conflict in intersection condition has been built.

The experimental results show that the model can well simulate the actual traffic conditions and improves the precision and practicability of the original model. But there still exists some idealized situations in the driving condition. At the same time, traffic data analysis or mining technology in this paper is not mature, and a large number of statistics and deep analysis should be done to improve the practicability and accuracy of the model.

References

1. Guo, Weiwei, Shaowei Qu, and Dianhai Wang. 2011. Traffic conflict discrimination model. *Journal of Jilin University (Engineering and Technology Edition)* 41 (1): 35–40.
2. Wang, Nan. 2014. *Research of the traffic conflict model based on the critical conflict area*. Beijing: North China University of Technology.
3. Feng, Guangdong, Wuhong Wang, Jianshuai Feng, and Huachun Tan. 2010. Modelling and simulation for safe following distance based on vehicle braking process. *IEEE International Conference on E-Business Engineering* 66: 385–388.

Research on the Vehicle–Bicycle Conflict Model at Signalized Intersection

Yu-quan Wang, Fang Xing and Liang Zhang

Abstract A model of vehicle–bicycle conflict at signalized intersections based on different vehicles was carried out by analyzing microscopic model to study the vehicle–bicycle conflict quantitatively. Traffic flows of different moving targets were taken into account in the research to analyze the conflict types of vehicles and bicycles. A mathematical model of vehicle–bicycle conflict at signalized intersections was established according to the location, speed, acceleration, and other parameters of the moving targets. Finally, taking the intersection between Haidian District Bajiao East Street and Era Garden North Road for instance, it is considered to be a conflict when the actual distance between the moving targets is less than the critical conflict radius, and there is no conflict otherwise. The results of the study demonstrate that the mixed traffic conflict model fits well with the actual situation. Besides, it is helpful to alleviate the traffic conflict and to realize the traffic safety at signalized intersection.

Keywords Traffic safety · Vehicle–bicycle conflicts · Critical conflict zone · Intersection

Introduction

The bicycle traffic takes the uppermost proportion of the city passenger transport in our country at one time in the past. With the increasing ownership of vehicle in China, urban transportation system is becoming more and more complex. As the bottleneck of urban road network, intersections bring together mixed and conflicting traffic flows which are in different ways and peculiarities such as motor vehicles, non-motor vehicles, and pedestrians [1]. Therefore, the serious problem between vehicle and bicycle has been occurred which will result in the reduction in

Y. Wang (✉) · F. Xing · L. Zhang
Beijing Key Lab of Urban Road Traffic Intelligent Tech,
North China University of Technology, Beijing 100144, China
e-mail: wyq@ncut.edu.cn

traffic capacity of the intersection. In reference [2], it analyzes vehicle–bicycle conflict at intersection. The effect of the conflict on the capacity of intersection is also researched. In reference [3], the paper introduces the method of gathering data of traffic conflict in detail on the basis of analyzing the present research situation of the traffic conflict at home and abroad. Then, some useful data such as the distance and the speed of the vehicle when the conflict occurs are obtained through video recording observation and the time to collision is also gained. The paper also does some research on the severity of the conflict at highway intersections according to the cumulative percentage of the time to collision. In reference [4], the concept of critical conflict zone is proposed from microscopic view, and the mixed traffic conflict discrimination model is established by taking the speed, distance, and angle as key variables.

At present, the research on traffic conflict is mostly between motor vehicles, and the research on traffic conflict of motor vehicle and non-motor vehicle is relatively little, so the problem of vehicle-bicycle conflict has not been solved yet [5], and it also impacts the control effect of the urban road traffic. Therefore, this paper classifies the traffic conflict at the intersection and establishes a mathematical model of vehicle–bicycle conflict at intersections according to the location, speed, acceleration, and other parameters of the moving targets. It facilitates to evaluate the safety of the intersection in a more accurate way, to ensure the traffic safety of the intersection, and to improve the traffic efficiency of the intersection.

Analysis of the Traffic Conflict Type

Signalized intersection is a traffic resource that can realize the transformation of the traffic flow direction. However, conflicts between traffic flows from different direction will cause the traffic resource competition of the intersection. Thus, this kind of traffic conflict leads to the traffic chaos and traffic accident of the intersection [5]. Theoretically, in order to avoid conflict, traffic flows of different directions are separated from time by adding phases. However, with the increasing of the phase, the cycle of the signal lamp and the delay time will increase. The lesser the phase number is, the higher the probability of traffic conflict will be. Therefore, this paper takes the two-phase and three-phase as examples to analyze the type of traffic conflict.

Two-Phase Signalized Intersection

There are two typical vehicle–bicycle conflicts at two-phase signalized intersection. One is the conflict between left-turn bicycle and different direction vehicles [6, 7], and the other is the conflict between straight bicycle and different direction vehicles. When a bicycle intends to turn left, the turning radius is larger than vehicle and its

Fig. 1 Main vehicle–bicycle conflict at two-phase signalized intersection
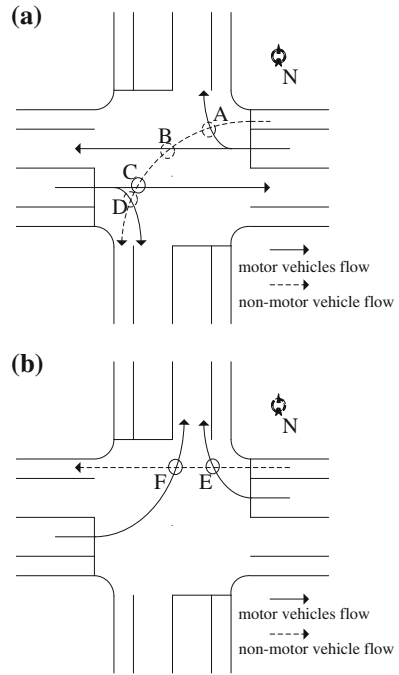
speed is lower than vehicle, which cause a great possibility of vehicle–bicycle conflict [8]. Besides, when a bicycle goes straight through the intersection, there are also conflicts with the right-turn motor vehicle in the same direction and the left-turn motor vehicle in the opposite direction.

As shown in Fig. 1a, taking the left-turn non-motor vehicles at the east imported as an example: Conflict A is the conflict between left-turn non-motor vehicle and the right-turn motor vehicle in the same direction; conflict B is the conflict between left-turn non-motor vehicle and the straight motor vehicle in the same direction; conflict C is the conflict between left-turn non-motor vehicle and the straight motor vehicle in the opposite direction; and conflict D is the conflict between left-turn non-motor vehicle and the right-turn motor vehicle in the opposite direction. As shown in Fig. 1b, taking the straight non-motor vehicles at the east imported as an example: Conflict E is the conflict between straight non-motor vehicle and the right-turn motor vehicle in the same direction [9]; conflict F is the conflict between straight non-motor vehicle and the left-turn motor vehicle in the opposite direction.

Three-Phase Signalized Intersection

The vehicle–bicycle conflict at auxiliary road of three-phase signalized intersection is similar to two-phase signalized intersection. Since the main road has taken

Fig. 2 Main vehicle–bicycle conflict at three-phase signalized intersection
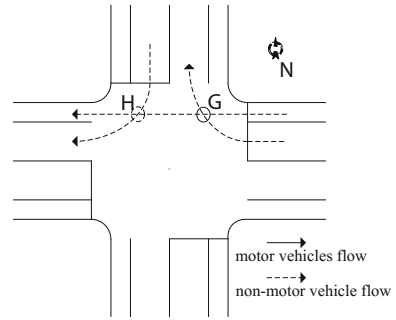
exclusive left-turn traffic signal lights, the vehicle–bicycle conflict at main road of three-phase signalized intersection is mainly between straight non-motor vehicle and the right-turn motor vehicle [5]. As shown in Fig. 2, taking the straight non-motor vehicles at the east imported as an example: Conflict G is the conflict between straight non-motor vehicle and the right-turn motor vehicle in the same direction; conflict H is the conflict between straight non-motor vehicle and the right-turn motor vehicle in the side direction.

In summary, the conflict between non-motor vehicle and motor vehicle at the signalized intersection is divided into the following categories:

a. the conflict between non-motor vehicle and the right-turn motor vehicle in the same direction, such as conflict A, conflict E, and conflict G;
b. the conflict between left-turn non-motor vehicle and the straight motor vehicle in the same direction, such as conflict B;
c. the conflict between left-turn non-motor vehicle and the straight motor vehicle in the opposite direction, such as conflict C;
d. the conflict between left-turn non-motor vehicle and the right-turn motor vehicle in the opposite direction, such as conflict D;
e. the conflict between straight non-motor vehicle and the left-turn motor vehicle in the opposite direction, such as conflict F;
f. the conflict between straight non-motor vehicle and the right-turn motor vehicle in the side direction, such as conflict H.

Establishment of Conflict Model

At present, methods of conflict discrimination have been proposed such as spatial distance method, time distance method, energy discrimination method, acceleration discrimination method, and conflict field theory. However, none of these methods are quite suitable for the microscopic movement of motor vehicle and non-motor vehicle. Therefore, this paper adopts the method of critical conflict zones. (There are certain sensitive areas around each traffic entities. When the other entities get

Fig. 3 Critical conflict zone model of vehicle–bicycle conflict
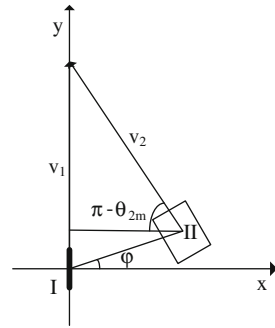
into the region and gradually draw near, the driver's mental stress will cause the two drivers to take reduction or steering and other hedging behavior [10]. It is considered to be a traffic conflict, and the sensitive area is the critical conflict zone [4].) In this paper, the motion trajectory and the motion law of the object are analyzed. Then, the model based on critical conflict zone is established.

In Fig. 3, I is a bicycle, II is a motor vehicle. We establish rectangular coordinate system, by taking the bicycle as origin of coordinates and the direction of its speed as y-axis. Assuming the length of the bicycle is l_1, the speed is v_1, the reduction rate is a_1, and the length of the vehicle is l_2, the width is d, the speed is v_2, the reduction rate is a_2, the angle between initial position of the motor vehicle and x-axis is φ. When the conflict occurs, the angle between the motion direction of the vehicle and the x-axis is θ_{2m}. Obviously, critical conflict zone radius R is a function of v_1, v_2, φ, and θ_{2m}, and θ_{2m} is determined by v_1, v_2, and φ. Therefore, R is a function of v_1, v_2, and φ [9], that is,

$$R(\varphi) = f(v_1, v_2, \varphi).$$

Assuming that the speed component of v_2 relative to the v_1 is v_{21}, so $|v_{21}|\cos\varphi = |v_2|\cos(\pi - \theta_{2m})$, that is [11],

$$|v_{21}| = |v_2|\frac{\cos\left[\varphi - \arcsin\left(\frac{|v_1|}{|v_2|}\cos\varphi\right)\right]}{\cos\varphi} \tag{1}$$

In the same way, $|a_{21}|\cos\varphi = |a_2|\cos(\pi - \theta_{2m})$, that is,

$$|a_{21}| = |a_2|\frac{\cos\left[\varphi - \arcsin\left(\frac{|v_1|}{|v_2|}\cos\varphi\right)\right]}{\cos\varphi} \tag{2}$$

Generally, the vehicle braking time can be divided into three periods: the driver's reaction time, the time from the driver stepping on the brake pedal to the vehicle braking, and the time of the vehicle braking force enhancement. However, due to the conflict can be observed only at the stage of generating the braking force,

the braking distance of the vehicle where the driver's reaction time is not included in the traveling distance [12], that is:

$$S = \frac{v^2}{2a} + v\left(t_1 + \frac{1}{2}t_2\right) \tag{3}$$

where: t_1 is the time form the driver stepping on the brake pedal to the vehicle braking; t_2 is the time of the vehicle braking force enhancement`.

Bringing the v_{21} and a_{21} of the above analysis into the formula (3), the critical conflict radius of vehicle–bicycle conflict in [0, 2 pi) can be obtained by considering the maximum component of the vehicle. (The bicycle is half of its length, and the vehicle is half of its diagonal length.)

$$R(\varphi) = |v_2| \frac{\cos\left[\varphi - \arcsin\left(\frac{|v_1|}{|v_2|}\cos\varphi\right)\right]}{\cos\varphi}\left(t_1 + \frac{1}{2}t_2 + \frac{|v_2|}{2|a_2|}\right) + \frac{1}{2}\left(l_1 + \sqrt{l_2^2 + d^2}\right) \tag{4}$$

For our aspect reading, assuming the length of the bicycle is 1.6 m, the speed is 10 km/h, and the speed of the vehicle is 15 km/h, the angle between initial position of the motor vehicle and x-axis is 60°, the mean value of t_1 is 0.09 s. The length of the vehicle, the width of the vehicle, t_2, and deceleration can be viewed in Table 1 [12]. So, the critical conflict radius of vehicle–bicycle conflict can be obtained as shown in Table 2.

Table 1 Maximum braking deceleration and braking time of various vehicles

Vehicle type	Maximum braking deceleration a (m/s^2)	Time of the vehicle braking force enhancement t_2(s)
Compact car	7.4	0.24
Medium-size car	6.2	0.36
Full-size bus	5.5	0.47

Table 2 Critical conflict radius of various vehicles at vehicle–bicycle conflict

Vehicle type	Vehicle length l_2 (m)	Vehicle width d(m)	Representative style	Critical conflict radius R (m)
Mini car	3.7	1.4	Chery QQ	5.9
Medium car	4.5	1.7	Beijing Hyun	7.1
Medium jeep	4.8	1.9	Jeep Cherokee	7.3
Full-size bus	12	2.5	Beijing JingHua	11.4

There are accuracy rate and conflict detection rate for evaluating the efficiency of vehicle–bicycle conflict model at signalized intersection. The accuracy rate is the ratio between the number of correctly detected conflicts and the total number of traffic conflicts detected by the discriminant model; the conflict detection rate is the ratio between the number of correctly detected conflicts and the number of the total conflicts which actually occurred [13].

Analysis of the Actual Examination

To verify the model, the intersection between Haidian District East Barkhor Street and Era Garden North Road is chosen. There are the five-minute video from three periods and 72 traffic entities in total. The position, speed, acceleration, and other information of the moving objects are extracted. Then, the critical conflict radius is calculated according to the intersection vehicle–bicycle conflict model. If the actual distance between the moving targets is less than the critical conflict radius, it is considered to be a conflict, and there is no conflict otherwise. Finally, the conclusion is compared with the actual motion of the moving object, and the model is validated.

In Fig. 4, I is a left-turn non-motor vehicle of South import, II is a left-turn motor vehicle of South import. We establish rectangular coordinate system, taking the bicycle as origin of coordinates and the direction of its speed as y-axis. Besides, the length of the bicycle is 1.7 m, the speed is 7 km/h, and motor vehicle is Beijing Hyundai, body size is 4.570*1.775 (m), the reduction rate is 2 m/s^2, and the speed is 10 km/h, the angle between initial position of the motor vehicle and x-axis is $-140°$.Then, bringing them into the formula (4), the critical conflict radius R is calculated as 4.36 m. The actual distance between bicycle and motor vehicle is about 3 m, which is smaller than the critical conflict radius, so there is a conflict. In the video, there is a conflict. And the result shows that the model is correct.

Overall speaking, 57 conflicts were detected in three videos according to the artificial detection. A total of 54 conflicts were discriminated by the vehicle–bicycle

Fig. 4 An example of vehicle–bicycle conflict

conflict model, which included the 45 proven conflicts. It can be concluded that the accuracy rate of conflict determination is 83.33% and the detection rate is 78.95%.

Conclusions

The conflict of motor vehicles and non-motor vehicles is an important factor to affect the traffic capacity of the intersection. It is helpful to improve the traffic capacity and the traffic safety of the intersection through the analysis of the intersection vehicle–bicycle conflict model. The paper analyzes the conflict types of vehicles and bicycles from microscopic view and establishes a mathematical model of vehicle–bicycle conflict at intersections according to the location, speed, acceleration, and other parameters of the moving targets. Finally, it is considered that the intersection vehicle–bicycle conflict model is consistent with the actual situation. Besides, it has certain practical values to solve the conflict problem in the mixed traffic environment.

References

1. Shen, J.J., W. Wang, and X.W. Chen. 2010. Study on conflict probability of motor and non-motor mixed traffic at urban intersections. *Journal of Southeast University (Natural Science)* 5: 1093–1096.
2. Jing, C. G., and D. H. Wang. 2004. Analysis and dealing method to conflict of mixed traffic at typical intersection. *China Civil Engineering Journal* 37 (6): 97–100.
3. Lu, C., Q.J. Xiang, G.Q. Zhang, et al. 2008. Determination of the severity of traffic conflict at highway intersections[J]. *Journal of Hefei University of Technology (Natural Science)* 31 (5): 683–686.
4. Guo, W.W., Z.W. Qu, and D.H. Wang. 2011. Traffic conflict discrimination model[J]. *Journal of Jilin University (Engineering Science)* 41 (1): 35–40.
5. van der Horsta, R.A., M. de Goedea, S. de Hair-Buijssenb, and R. Methorstc. 2014. Traffic conflicts on bicycle paths: A systematic observation of behaviour from video. *Accident Analysis and Prevention* 62: 358–368.
6. Sun, F. 2010. *Operating left-turn phase at signalized intersection based on traffic conflict technique. The 10th International Conference of Chinese Transportation Professionals*, 556–564. Beijing: American Society of Civil Engineers.
7. Xu, L.J., and W. Wang. 2006. Analysis of Influence of Left-Turn Non-motors in Signalized Intersection. *China Highway Journal* 1 (19): 89–92.
8. Wang, N. 2014. *Research of the traffic conflict model based on the critical conflict area[D]*. Beijing: North China University of Technology.
9. Autey, J., Sayed, T., and M.H. Zaki. 2011. Safety evaluation of right-turn smart channels using automated traffic conflict analysis. *Accident Analysis and Prevention* 45: 120–130.
10. Shao, H.J. 2008. *Study on the mixed traffic conflict probability at urban signalized intersections[D]*. Beijing: Beijing Jiaotong University.
11. Older, S.J., and ShiPPy J. 1979. Proceedings of the Second International Traffic Conflict Workshop. 6(5): 21–25.

12. Liu, X.M., and H.L. Duan. 1997. Research on Standard Program of Traffic Conflict Techniques at Intersections. *Highway Traffic Science and Technology* 14 (3): 29–34.
13. Zeng, Y.S., Ma, S.F., Zhong, S.Q. et al. 2012. A video detection-based traffic conflict classification model[J]. *Journal of Transport Information and Safety* 4: 11–14.

Research of Variable Lane Control Method in the Emergency Evacuation Area

Lili Zheng, Xinyue Hu and Tongqiang Ding

Abstract When unexpected disasters are coming, tens of thousands of people lost their lives because of no emergency aid caused by traffic jam. We need to build an emergency planning based on traffic evacuation. This paper proposes a method of variable lane control based on emergency evacuation area, establishes a bi-level programming model and designs solution algorithm based on harmony search to find the optimal solution quickly. After getting the optimal solution, we build the simulation experiment to verify the feasibility and effectiveness of proposed method according to the comparison of the saturation degree and the total travel time of system.

Keywords Emergency evacuation area · Method of variable lane control · Bi-level programming model · Harmony search algorithm

Introduction

Since the twentieth century, natural disasters in worldwide have been increasing frequently from the earthquake to the tsunami and the hurricane. People have studied urgently with all kinds of typical emergency plans focused on traffic evacuation.

L. Zheng
State Key Lab of Automobile Simulation and Control, Jilin University,
Changchun 130022, China

L. Zheng · X. Hu
College of Transportation, Jilin University, Changchun 130022, China

L. Zheng · T. Ding (✉)
Jilin Province Key Laboratory of Road Traffic, Jilin University,
Changchun 130022, China
e-mail: dtq8@163.com

Variable lane control which means the variation of the direction of one or more lanes changes all or parts of the lane directions from driving to the affected areas for driving away the affected areas, which can improve the evacuation capability of traffic network greatly and reduce the evacuation time. In 1998, when Florida and Georgia suffered the hurricane, the local administration of America used the variable lane strategy for single interstate road; consequently, the effect is remarkable. In 2005, under the condition of dealing with the evacuation about 3,000,000 people during the hurricane "Rita" in the USA, when the serious congestion occurred (heavy traffic moved 10–20 miles in nine hours), local government implemented variable lane solutions to the interstate 45, interstate highway 290, and interstate 10. The solution solved the jam partly of interstate highway 290, but after a few hours, this solution was forced to be annulled because of the traffic jam around interstate highway 290 [1]. Qiang Sun researched and developed the optimal allocation principles to make a single lane road to adapt to the asymmetry of traffic demand in one day. The capacity of traffic network was improved by the establishment of the most optimized model and the design of the step of dynamic variable lanes [2]. During the process of the practice, even though the fact that reversing local lanes direction could increase the evacuation capacity of part of the road, it could lead to the appearing of traffic bottleneck point in other lanes because of lack of consideration as a whole and reduce the evacuation performance of the whole traffic network. This paper proposes a method of variable lane control based on traffic network, establishes a bi-level programming model, and designs the harmony search to solve the model [3].

The Establishment of the Model of Controlling the Variable Lane Based on Traffic Network

From the point view of the traffic managers, when they adjust the lane, they hope to alleviate traffic jams, named the minimal traffic impedance. From the point view of traffic participants, they choose the shortest path by instinct, named Wardrop balanced principle. Hence, this paper considers the bi-level programming model about the adjustment of variable lane from two aspects [4]. Firstly, in the view of system, the establishment of the upper planning model introduces the minimum of the total impedance of the system. Secondly, in the view of traffic participants, the establishment of the lower-level planning model demands to be applied to the user equilibrium model. After getting each section of the optimal flow distribution of every link, the result will be used in the upper planning model and get the most appropriate lane volume of different directions and the adjustment scheme of variable lane [5].

Considering an urban traffic network consisted of nod-set N and two-way link-net A, each of the two-way traffic links is composed of two sections in the opposite directions [6].

where:

A	Two-way link-net;
a	Any link in the network;
x_a	Flow of any link a;
f_p^{rs}	Flow on path p between r and s;
q^{rs}	Traffic demand between r and s;
n_a	The number of both-way link, $n_a > 0$ as a integer;
k_a	Traffic capacity of single lane on link a;
u_a	Lane number of positive direction in link a after adjustment, value it $[0, n_a]$ as an integer;
δ_{ap}^{rs}	If link a is on the path p between r and s, it act as 1; or 0; and
$t_a(x_a, u_a)$	impedance function of link a.

The function adopts to the famous BPR:

$$t_a(x_a, u_a) = t_0 + 0.15 \times t_0 (x_a / u_a s_a)^4 \tag{1}$$

In order to minimize the total impedance of the whole network system,the department of city traffic management determines the variable lane adjustment scheme according to the situation of the traffic flow on the link. The traffic participants follow a path of least impedance according to the condition of road. The behavior of choosing the path conforms to the Wardrop balanced principle. All above can be described as the bi-level programming model as following:

The upper model:

$$\min_u Z = \sum_{a \in A} t_a(x_a, u_a) \cdot x_a \tag{2}$$

$$\text{s.t. } u_a \in \{0, 1, 2, \ldots, n_a - 2, n_a - 1, n_a\}, \forall\, a \in A \tag{3}$$

The lower-level model:

$$\min_x \sum_{a \in A} \int_0^{x_a} t_a(w, u_a) dw \tag{4}$$

$$\text{s.t. } \sum_p f_p^{rs} = q_{rs}, \forall\, r, s \tag{5}$$

$$x_a = \sum_r \sum_s \sum_p f_p^{rs} \delta_{ap}^{rs}, \forall\, a \in A \tag{6}$$

$$x_a \geq 0, \forall\, a \in A \tag{7}$$

Algorithm Based on Harmony Search

Harmony search algorithm is a new type of the method of the intelligent optimization introduced in recent years with the advantages of small time complexity, wide applicable scope, and simple structure and operation. The arithmetic can be used in both linear integer-programming problem and nonlinear integer-programming problem and have a rather practical value on the application layer for solving NP-hard problem. It can obtain the same or even better solution, compared with the genetic simulated annealing algorithm and hybrid genetic algorithm [7]. The theory imitates the process of playing music. The optimal effect of harmony is attained, according to a few primitive harmony to adjust again and again. Each kind of instruments in band is a variate in the objective function, and tone is the value of variate in the objective function [8]. Same as the musical performance, optimized algorithm is to looking for various optimal states determined by the objective function values such as minimum cost, maximum benefit, or the highest efficiency.

The model in this paper is nonlinear mixed integer bi-level programming problem, which is widely recognized as one of the optimization problems that is extremely difficult to solve. The upper decision variates are integers; the lower decision variates are real numbers. This paper is based on the harmony search algorithm to solve the upper planning and based on the Frank–Wolfe method to design the flow distribution in the lower planning. The following is its basic process:

Step1: Initialize the optimized problem and algorithm parameter. Set the size of harmony memory (HMS) and maximum iterations (N_{max}). Number of HMS should be much smaller than all the feasible solutions.
Step2: Initialize the harmony memory, determine the scope of $u_a[0, n_a]$, and generate HMS solutions of optimized problem randomly, then put these in harmony memory, which expresses as following:

$$\left\{ \begin{bmatrix} u_1^1 & \cdots & u_i^1 & \cdots & u_n^1 \\ & \vdots & & \ddots & & \vdots \\ u_1^m & \cdots & u_i^m & \cdots & u_n^m \\ & \vdots & & \ddots & & \vdots \\ u_1^{HMS} & \cdots & u_i^{HMS} & \cdots & u_n^{HMS} \end{bmatrix} \middle| \begin{array}{c} Z(x^1) \\ \vdots \\ Z(x^m) \\ \vdots \\ Z(x^{HMS}) \end{array} \right\}$$

Among them, $U^m = (u_1^m, u_2^m, \ldots, u_n^m)$ is the mth solution vectors. The corresponding objective function is $Z(x^m)$.

Step3: Generating a new solution. There will be a new solution created $U^{\text{new}} = (u_1^{\text{new}}, \ldots, u_i^{\text{new}}, \ldots, u_n^{\text{new}})$. Among them, the u_i^{new} will be produced by the following three ways:

1. Keep some components in the harmony memory.
2. Generate new components randomly.
3. Disturb the solution generated in the above method.

Keeping some components in the harmony memory and keeping some components in the harmony memory at a certain probability mean the probability that new solution u_i^{new} is from set of ith $U_i = (u_i^1, u_i^2, \ldots, u_i^{\text{HMS}})^{\text{T}}$ in the harmony memory is HMCR. Generate randomly means that the probability that the new solution is from the outside of the feasible solution space in harmony memory is 1-HMCR. Disturb the solution components retained from the above steps u_i^{new} in certain probability (PAR), working on following principle:

$$u_i^{\text{new}'} = u_i^{\text{new}} + 2 \times u \times \text{rand} - u$$

where:

u_i^{new}	The solution before the disturbance;
$u_i^{\text{new}'}$	The solution after the disturbance;
u	Bandwidth; and
r and	A random number valued 0 or 1

$$u = bw \times (gn/\text{max})$$

The value of bw is between 0.4 and 0.6. gn is the rest of iterative times. N_{max} is the maximal iterative times. u is rounded down.

Step4: Update the harmony memory. If the new is better than the worst memory, make the new solution replace the worst one in HMS.
Step5: Distribute the traffic flow in user equilibrium.
Step6: If the outcome fulfills the termination criterion, and n meets the maximal iterative times N_{max}, the optimal solution u_a^* outputs. If not, turn to step3.

Algorithm process is as follows (Fig. 1):

Simulation Verify

Simulation Data Instruction

The road network topology adopted in this paper is as shown in Fig. 2 and the road link data is as shown in Table 1 [9].
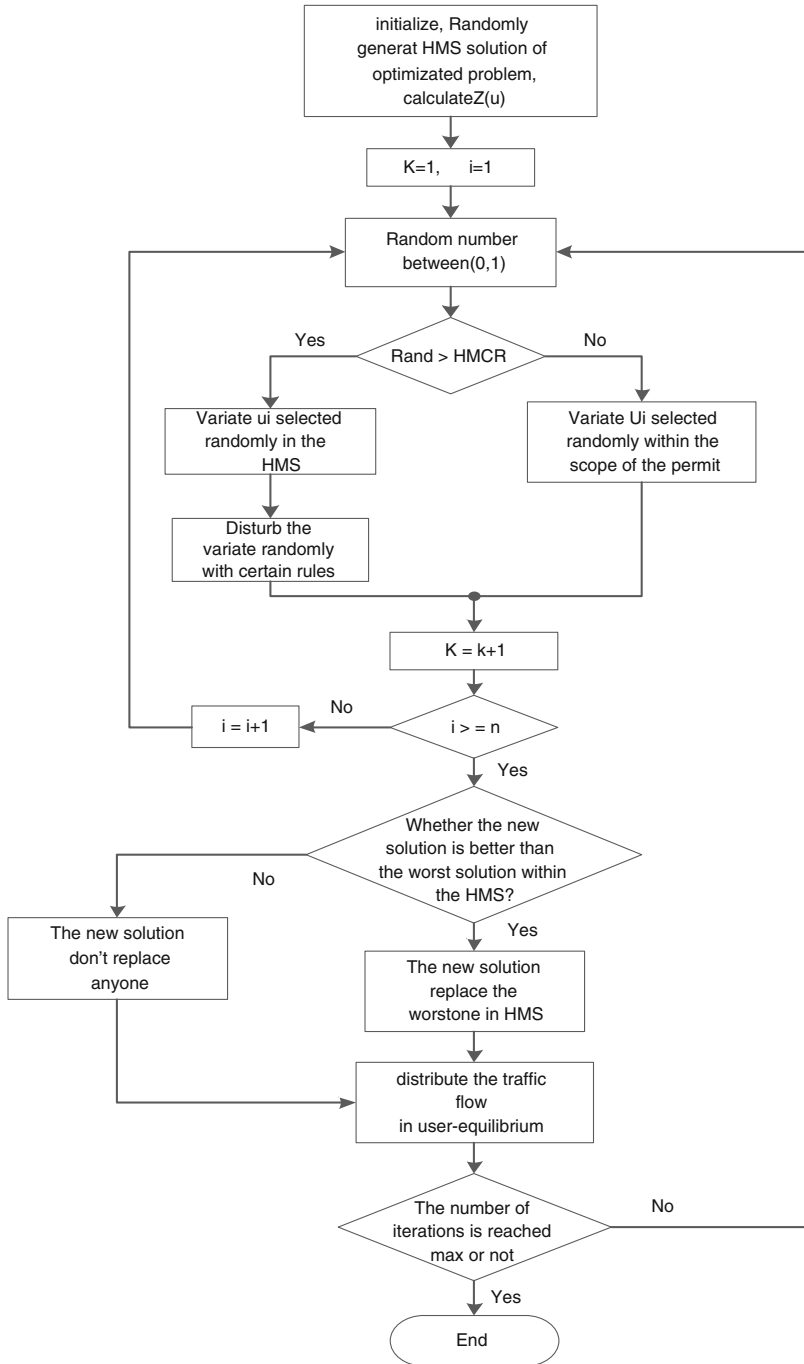
◄Fig. 1 Algorithm process about the V lane control model in emergency evacuation area based on harmony search

Fig. 2 Road network topology
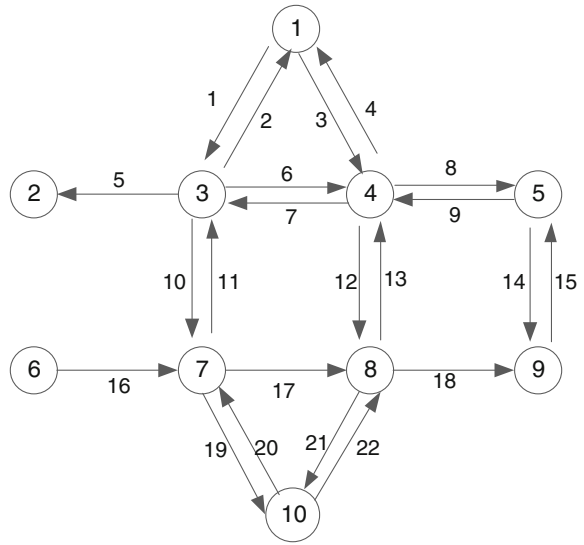
Table 1 Link data

Link	t_0 unit: s	k_a unit: pcu/h	n_a
1, 2	200	1800	8
3, 4	200	1800	4
5	200	1800	2
6, 7	200	1800	2
8, 9	200	1800	2
10, 11	260	1800	8
12, 13	200	1800	8
14, 15	200	1800	2
16	200	1800	2
17	200	1800	2
18	200	1800	2
19, 20	200	1800	8
21, 22	200	1800	8

There are four couples of O–D: $5 \rightarrow 2$, $6 \rightarrow 9$, $1 \rightarrow 10$, and $10 \rightarrow 1$. Corresponding traffic volume is 1826, 468, 6610, and 1836 pcu/h.

Designing Simulation Environment

1. Build simulation environment.

In the road network, such as Fig. 2, there is a sudden traffic accident at the intersection No. 8, so that the link around the intersection No. 8 cannot work normally.

2. Describing contradistinctive plans.

Program0: Lane directions without any adjustment.
Program1: In the situation of designing simulation, direction of the traffic flow is distributed imbalanced in different direction between node1 and node10. The traffic flow on $1 \rightarrow 10$ is significantly higher than the traffic flow on $10 \rightarrow 1$. Under the above unexpected incidents, the traffic managers adjust the number of lanes artificially in order to avoid the extreme congestion of relevant roads. Adjustment scheme is that link20, link11, and link2 lend a lane to the different direction link, which is link1, link10, and link19;
Program2: Determine the adjustment scheme using the method in this paper.

3. Contrast indexes

The paper selects the saturation of each section and total travel time as contradictive indexes and compares three types of schemes, respectively, from the angle of local and system, in order to evaluate the service level of the road network from different schemes.

Analyze the Outcome of Comparison

Different adjustment schemes are as follows (Tables 2 and 3):

1. The saturation of each section (Table 4)

The data attained from the VISSIM is used to calculate the ratios of each section, as shown in Fig. 3.

Program0: Without any adjustment. Individual sections have been serious congestion, traffic flow is quite unstable.
Program1: After the artificial adjustment, the ratios of some sections have dropped slightly, but the whole ratios remain high. The network is unstable;
Program2: Under the condition of program 2, the whole ratio of network has obvious downward trend and every ratio is between 0.1 and 0.63. The network is in a stable state, according to the evaluation indexes of the HCM.

2. The contrast between total travel time.

Concluding from the histogram as follows, the total impedance of system is the minimum one in program 2 compared with other two schemes. Compared with program1, program2 has better optimization effect (Fig. 4).

Table 2 Adjustment scheme in Program1

Link	1	2	3	4	5	6	7	8	9	10	11
No. of lane	5	3	4	4	2	1	3	1	3	5	3
Link	12	13	14	15	16	17	18	19	20	21	22
No. of lane	4	4	2	2	2	2	2	5	3	4	4

Table 3 Adjustment scheme in Program2

Link	1	2	3	4	5	6	7	8	9	10	11
No. of lane	6	2	4	4	2	1	3	1	3	6	2
Link	12	13	14	15	16	17	18	19	20	21	22
No. of lane	4	4	2	2	2	2	2	6	2	4	4

Table 4 Index of service quality evaluation

Reference value of service quality evaluation (traffic volume/capacity)	
V/C	The characteristics of operation
≤ 0.35	Free-running traffic flow (smooth)
0.35–0.55	Reasonable free-running traffic flow (a little bit delay)
0.55–0.75	Stable traffic flow (receptible delay)
0.75–0.90	Approach instable traffic flow (enduring delay)
0.90–1.00	Unstable traffic flow (crowded)
>1.00	Traffic jam
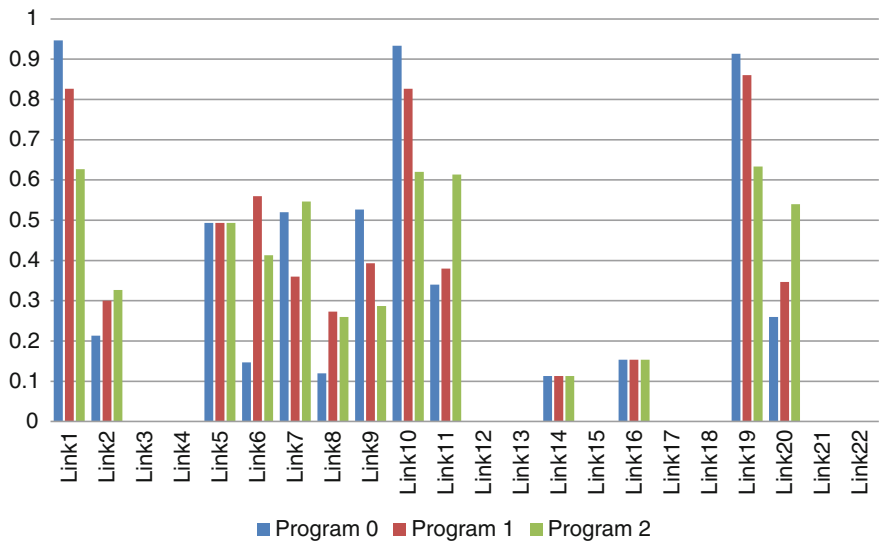
Fig. 3 Contrast figure of saturation in three kinds of solution

In conclusion, the adjustment scheme proposed in this paper reduces the extent of the road congestion and the impedance of the system and makes full use of the resources of the road, which improves the capacity of the network enormously and greatly.

Fig. 4 Comparison about travel time in different schemes
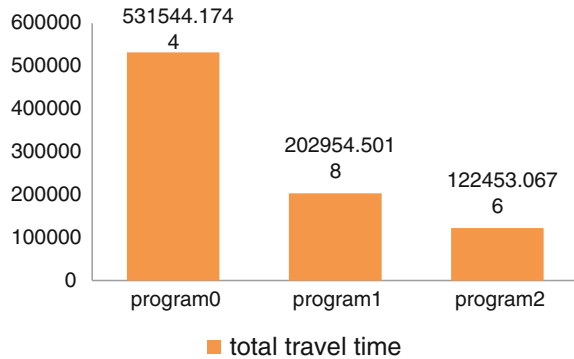
Conclusion

The paper studies the methods of the variable lane controlling which are applied to the emergency evacuation area, then establishes a bi-level programming model—the upper plan that aims at the minimum total travel time and the lower-level plan described by user's equilibrium principle—and designs the solution algorithm based on the harmony search, finally, verifies the proposed method by simulation. Conclusions show that the proposed method is feasible and effective in the traffic emergency evacuation.

Aimed at large-scale traffic network, the author will research the corresponding parallel algorithm in the following study in order to improve the operation efficiency of the algorithm.

References

1. Kim, Sangho, Shashi Shekhar, and Manki Min. 2008. Contraflow transportation network reconfiguration for evacuation route planning. *Transactions on Knowledge and Data Engineering* 20 (8): 1115–1129.
2. Sun, Qiang. 2009. Research about algorithm of dynamic variable lane optimization. *HIGHWAY* 8.
3. Suwansirikul, C., T.L. Friesz, and R.L. To bin. 1987. Equilibrium decomposed optimization: A heuristic for the continuous equilibrium net-work design problem. *Transportation Science* 21.
4. Huang, Shuai, and Liang Ma. 2014. Solving general integer programming problem by improved harmony search algorithm. *Computer Engineering and Applications* 50 (3).
5. Hao, Wang, Liqun Gao, and Haibin Ouyang. 2014. A modified harmony search algorithm with global crossover. *Journal of Northeastern University (Natural Science)* 35 (8).
6. Zhang, Haozhi, and Ziyou Gao. 2007. Optimization approach for traffic road network design problem. *Chinese Journal of Management Science* 15 (2).

7. Wang, Xiong. 2013. *Bi-level programming model and algorithm of urban roads convertible lanes*. Changsha: Zhongnan University.
8. Gao, Ziyou, Haozhi Zhang, and Huijun Sun. 2004. Bi-level programming models, approaches and applications in urban transportation network design problems. *Journal of Transportation Systems Engineering and Information Technology* 2 (4).
9. Zhang, Peng, Wenquan Li, and Yulin Chang. 2010. Reserve capacity model for urban road network with variable lanes. *Journal of Southwest Jiaotong University* 45 (2).

Modeling the Traction Energy Consumption for Urban Rail Line Considering Operation Characteristics

Xun Sun, Zhankui Ma, Enjian Yao and Xu Wu

Abstract With the rapid construction of urban rail transit systems in China, the operating mileage and the passenger volume of urban rail transit system have been growth rapidly, as well as huge energy consumption has been generated. How to accurate evaluate and predict the energy consumption has attracted more and more attention of researchers, operating enterprise, and policy-makers. Therefore, this paper is devoted to establishing a traction energy consumption model to accurate evaluate and predict the traction energy consumption which is the main energy consumption in urban rail transit systems. On the basis of dynamics analysis of urban rail transit, four categories variables are adopted into the model, including train running state, line attribute, operation indexes and environment variable factors. Then, two multiple regression models of the traction energy consumption factor for underground line and overground line are proposed. Based on the urban rail transit statistics database of Beijing, the models are calibrated by the least square method. The results show a good-fit between the proposed model estimates and the field measurements, and the proposed model has satisfied prediction accuracy. Furthermore, to further improve the prediction accuracy, two extend models involving the season factor for underground line and overground line are proposed in this paper. The calibrated results show that the extended model is more effective to evaluate the traction energy consumption factor than the previous model, and the MAPEs of the extended model for underground line and overground is 4.7 and 5.1%.

X. Sun (✉) · Z. Ma · E. Yao
MOE Key Laboratory for Urban Transportation Complex Systems
Theory and Technology, Beijing Jiaotong University, Beijing, China
e-mail: xsun1@bjtu.edu.cn

Z. Ma
e-mail: 14125716@bjtu.edu.cn

E. Yao
e-mail: enjyao@bjtu.edu.cn

X. Sun · Z. Ma · E. Yao · X. Wu
School of Traffic and Transportation, Beijing Jiaotong University,
Haidian District, Beijing, China
e-mail: wuxu@bjtu.edu.cn

Introduction

With China's urbanization progress, the rapid augment in urban population as well as the growing vehicle ownership, a series of problems such as traffic congestion, environmental pollution and energy crisis become the great barrier to the sustainable development of China. Developing urban rail transit is regarded as a prior way to solve traffic problems due to its low energy consumption, low emission and low pollution. However, with the rapid development of urban rail transit, there exist lots of severe challenges. On a hand, urban rail transit plays a more important role in satisfying the urban resident travel demand and alleviating the urban traffic pressure greatly. Residents' increasing reliance on urban rail transit promotes the expansion of rail transit network, and meanwhile, it brings higher requirements to the transportation service level. On the other hand, with the growing energy consumption of urban rail transit every year, the rapid energy consumption is becoming a serious problem to be solved. As a result, energy conservation and emission reduction will be a key task in urban rail transit operation.

One of the severe problems of urban rail transit system is that the total energy consumption is huge, of which the traction energy consumption makes up the largest proportion, accounting for about half of the total energy consumption. Usually, the traction energy consumption is influenced by passenger volume, environment temperature and etc. These factors together determine the total amount of the traction energy consumption. Therefore, it is essential to analyze the factors which influence the traction energy consumption, understand the functional mechanism and impact of each factors. And then based on the factor analysis, forecasting the traction energy consumption can be utilized to provide the management basis, support operation decision-making, acquire the efficiency of urban rail transit energy consumption, and provide guidance to energy conservation and emission reduction in rail transit operation.

Literature Review

At present, the studies on energy consumption model train track are mainly conventional rail and high-speed railway train traction energy consumption. However, the urban rail transit system has its own characteristics. For example, it has short station spacing, starts and stops frequently, has a relatively low speed. So its energy consumption model should be discussed according to its own characteristics.

UMT Traction energy consumption is an important part of the energy consumption of rail operations. Danziger [1] pointed out that energy consumption of rail transport is a complex system, and the energy consumption analysis should consider the importance of the various systems, identify the key factors that affect energy consumption. In this complex urban rail transit system, the factors that influence the energy consumption of train operation are complex and diverse, including the weight of the train, the characteristics of train traction, the slope of the line, the line turning radius, the average station spacing, speed, group program, stop program, ambient temperature, passenger turnover, distance traveled, load factor and so on. Hoyt and Levary [2] believed that the main factors affecting the traction energy are trains property, geographical attributes, and other uncertain factors such as wind speed, precipitation, temperature and other. Important factors such as load, running speed and number of train stations which affect traction energy consumption was analyzed by Hickman et al. [3]. Based on the analysis on the factors affecting energy consumption of urban rail transit, Wu et al. [4] constructed a model of which the unit passenger volume is the index and the passenger traffic, target speed are the influencing factors. Peng et al. [5] analyzed the influence on railway goods transportation energy consumption from the aspects of infrastructure, mobile devices and operation organization. Liu et al. [6] analyzed the line's curves, ramps and train weight which may influence the energy consumption quantitatively. Xu et al. [7] believed that the factors that influence energy consumption are passenger capacity and the speed of the train. Through the analysis of influencing factors, Liu and Tian [8] established a kinematic equation whose variables are travel distance, passenger turnover and ambient temperature.

It can be seen that the analyses of influence factors are various and incomplete. There are generally two methods to analyze the factors that influence the energy consumption of rail transit traction. One is based on the analysis of the simulation model, which is a more experimental method and cannot be used in practice. The other one is based on the statistical analysis, which requires a lot of actual statistical data. The calculation of the data and the accuracy of the data are the major problems at present.

Fan et al. [9] analyzed the energy consumption at the station by establishing an improved cellular automata simulation model. Wu et al. [4] established a function conversion simulation model based on factors of air resistance, friction, climbing resistance. Albrecht [10] used the method of simulation experiment to verify the effectiveness of the optimal train running time. Wang [11] also used this method to analyze the sensitivity of the influencing factors of urban rail transit energy consumption from the train infrastructure attributes, line conditions and transport organization mode. Malavasi et al. [12] studied the influence of trains running on the Mass Rapid Transport energy consumption by building simulation models. Bocharnikov et al. [13] found the optimal driving strategy to reduce energy consumption by genetic algorithm traction. Gu et al. [14] proposed train energy-saving operation model which was based on real-time traffic information, and make the train energy policy based on the results of numerical analysis to save energy from the perspective of optimizing train running. Feng [15] analyzed

factors influencing rail transportation energy consumption based on statistics of the average passenger volume, speed technology and the monthly average temperature.

It can be seen that the studies based on statistics of rail transit energy consumption are still relatively scarce. We can do some research from this aspect. Most of the energy consumption estimation models estimate the rail transportation energy consumption based on kinematic equation, For example, Xu et al. [7] established a movement mechanics calculation model, with passenger traffic and the speed of the train as variables. Wang et al. [16] selected train weight, passenger capacity of traction energy consumption as influencing factors and calculate the energy consumption index value using motor type and laying method. Xue et al. [17] choose train quality, technology, speed, run mileage, gradient as the main influence factors of energy consumption and set up a kinematics equation to calculate the energy consumption. Wei et al. [18] assumed that urban rail trains' electromechanical efficiency is constant and then set up a kinematics equation to estimate train traction energy consumption, and also established an energy-saving operation model of urban rail transit. González–Franco et al. [19] believed high speed railway train energy consumption is mainly composed of resistance-overcoming energy, kinetic energy consumption, potential energy consumption and auxiliary traction energy consumption. Through detailed discussion of work of the traction in the process of train running, he established a traction power model to calculate the energy consumption.

In addition, there are some researches which calculate the energy consumption through the regression analysis model. Hickman et al. [3] proposed train traction energy estimation formula based on the average speed and station spacing. Liu and Tian [8] set up a multiple linear regression model, using the distance, passenger turnover, and environmental temperature as influence factors.

It's difficult for the calculation model based on the kinematic method to identify parameters. However, the calculation model based on statistical data may better fit the actual situation. Although regression calculation model is weak in portability, there are advantages in the parameter calibration and it can effectively predict the energy consumption according to changes in the relevant variables. Therefore, this article intends to analyze and estimate the affecting factors of urban rail transit energy consumption by regression analysis.

Energy Consumption Modeling

As one of the most important components in total operation energy consumption of urban rail transit system, the traction energy consumption of urban rail transit is influenced by various factors. In the paper, on the basis of dynamics analysis of urban rail transit, the key task is to analyze the major factors which influence the traction energy consumption.

Dynamics Analysis of Urban Rail Transit Train Operation

(1) Stress analysis of rail transit train

The traction energy consumption of rail transit system is the total energy consumption in the process of train operation. The energy working in the train movement process (pure traction energy consumption), which is the most important component of the total energy consumption, is basic energy consumption in order to implement the transportation service. In addition, there exists additional service energy consumption of air condition, ventilation, illumination and etc. in the train operation process. Therefore, the total energy consumption can be indicated as the main energy consumption above.

The train operating status can be divided into three parts: traction, coasting and braking. The resultant force on the train can be expressed in the three statuses above:

Traction:

$$C = F - W_0 \tag{1}$$

Coasting:

$$C = -W_0 \tag{2}$$

Braking:

$$C = -(W_0 + B) \tag{3}$$

where C represents the resultant force on the train; F represents train tractive force; W_0 represents train running resistance; B represents train braking force; c represents unit resultant force shown as below:

$$c = \frac{C \cdot 10^3}{(P + G)g} \tag{4}$$

where P represents mass of the train, t; G represents passenger mass, t.

From the above formulations, train tractive force, running resistance and braking force are mainly applied on the train of urban rail transit in the operating process.

(2) The traction and braking force

The traction energy consumption of rail transit train can be viewed as the energy working which is done by the traction force to overcome the resistance force with the guarantee of specific operation status and by the braking force to help deceleration. Therefore, the traction energy consumption of rail transit train can be viewed as the sum of work done by traction and braking force.

Traction force is the external force reacting at the wheel rim from the rail generated by the adhesion between wheels and rail, which is the internal force generated by locomotive power unit, and then form into the tangent direction of the force at the wheel rim by gearing. The train braking force is the external force to hinder the train operation generated by the braking device, which can be adjusted by the drivers. The train braking force is much stronger than the resistance force. And during the braking, although the resistance force is working as well, the train braking force develop the main effect in braking the train. Train velocity is the independent variable in the function of the train traction force and braking force, but the function is closely related to the resistance force and the driver's operation modes. It is difficult to estimate by the theory of simple formulation.

Based on the formula of kinetic energy, the formula of kinetic energy in the process of start and brake are shown as below:

Start phase:

$$W_f - W_0 = mv^2/2 \tag{5}$$

Brake phase:

$$-(W_c + W_0) = -mv^2/2 \tag{6}$$

From the formula above, the work done by the train traction force and braking force is related to the train running resistance and operating status. Therefore, the traction energy consumption of rail transit train can be estimated effectively on the basis of the train running resistance and operating status

(3) The train running resistance

The train running resistance is the external force reacting on the operating trains, yet not controlled by manpower. The basic train running resistance includes:

(1) Bearing resistance: related to the velocity and mass of train.
(2) Rolling resistance: related to the mass of train.
(3) Sliding resistance: related to the mass of train.
(4) Shock and vibration resistance: related to the quality of railway line and the velocity of train.
(5) Air resistance: directly proportional to the square of the velocity.

The resistances above exist in the process of train operation, whose formulas can be constructed with axle load, velocity and the square of the velocity.

Furthermore, there are some additional resistances, including

(1) Gradient resistance: related to the quality of railway line (gradient) and the mass of train.
(2) Curve resistance: related to the quality of railway line (curve length and curve rotation).

3) Air resistance for tunnel: related to the quality of railway line (tunnel length), train features (windward area and train length) and the velocity and mass of train.

Explanatory Variables

Through the study of the dynamics analysis of city railway train operation, it is not hard to find that traction energy consumption was under the influence of train running state, line attribute, operation indexes and environment variable factors.

Speed is the most important indicator of train running state, is the main affecting factor of city railway train traction energy consumption. All the resistance can be estimated by a function with the speed. In actual operation, the train needs to maintain the speed or acceleration acting traction force to overcome the resistance, which leads to an increase in energy consumption. With the increase of train speed, energy consumption also increases, the greater the speed the larger the magnitude of increase in energy consumption. Taking Guangzhou Metro as an example, the Guangzhou Metro Line 1, the train speed increases on average every 5 km/h, operation energy consumption increases by 13.1%. In Guangzhou Metro Line 3, train speed increases on average every 5 km/h, energy consumption increases by 10.3% [11].

Line attribute variables include the average station spacing, line length, line laying methods. Among them, the average station spacing directly affects the urban rail trains start and stop frequency, which may affect its energy consumption. In addition, according to the relevant analysis of the running resistance, it is not difficult to find the length of the line will also determine how many work trains running resistance do. Line laying way has the branch of the ground line and the underground line. Different line installation will have an effect on train traction energy consumption. This effect is mainly caused by air resistance and vehicle-mounted auxiliary facilities. Vehicle auxiliary equipment energy consumption is in the process of the train operation, in order to ensure the passengers' comfort and safety. When the line is the ground line, the air resistance is different from the tunnel underground line air resistance. Relatively, the underground line is less affected by environmental temperature.

When the vehicle and the passenger flow are under the same condition, the energy consumption of the underground line is greater than that of the ground line. The basic reason is that the trains in underground tunnel need to overcome more air resistance. Taking Beijing metro line 1 for example, in 2009 the average annual energy consumption per unit ton-km is about 268 kw h, and the same driving condition of subway BaTong line of this metric is 227 kw h. The former is about 18% [20] higher than the latter. Considering the different line laying way may influence on energy consumption, in the following research, the author will divide

the circuit into the ground line and the underground line. Data processing and the establishment of energy consumption calculation model are based on this.

In the operational indicators, the unit passenger volume together with the weight of the train determines the axle load of the train. Trials show that the resistance of locomotive, vehicle is proportional to the weight. So unit passenger volume and train weight are also important factors affecting train energy consumption.

When train running state is the same, but temperature is different, the city railway train traction energy consumption will also be different. As a result, the energy consumption model correction should be carried out using the environment temperature.

According to physical principle, the higher train running speed is, the greater the amount of electrical energy consumption will be. At the same time, the length of the line, the environment temperature will also affect the energy consumption of train. Moreover unit train passengers and train weight will affect the train axle load. Therefore, under the existing research, this paper uses the square of the speed, the reciprocal of the speed, speed, line length, average station spacing, passenger traffic, train weight, and the average transportation distance as the main influencing factors to build the energy consumption model, using ambient temperature as adjustment factor to modify the model, in order to establish a model which can describe the real orbit transportation energy consumption. And the correlation between the above variables with traction energy consumption is shown in Fig. 1.

Model Structure

At present, the research on the energy consumption model mainly has three kinds of models: the regression analysis model based on the data, the model based on the electric power and the model based on the kinematic method. However, the current energy consumption statistics of urban rail transit system are mainly statistics of the whole line for a period of time (including traffic, operating data). There is no microscopic dynamic data during the train operation process. Therefore, based on the above data, the regression analysis model is used to analyze the energy consumption model of urban rail transit train. In addition, considering the different ways of laying lines, in this paper, the subway lines are divided into two categories, which are the ground line and underground line, and the traction energy consumption model is constructed respectively.

In order to improve the applicability of the model to every line of the same type, the traction energy consumption in unit ton kilometers (hereafter referred to as traction energy consumption factor) is adopted in this paper as a measure of energy consumption in urban rail transit unit indicators. It is not difficult to find on the basis of dynamic analysis of urban rail train, traction energy factor consumption is related to the influence factors of train running state, line attribute, operation index and environment variable. Therefore, this paper establishes the multiple regression model of the traction energy consumption factor in urban rail transit as follows:

(a) Average speed

(b) Average station spacing

(c) Unit train passenger volume

(d) Unit train passenger volume
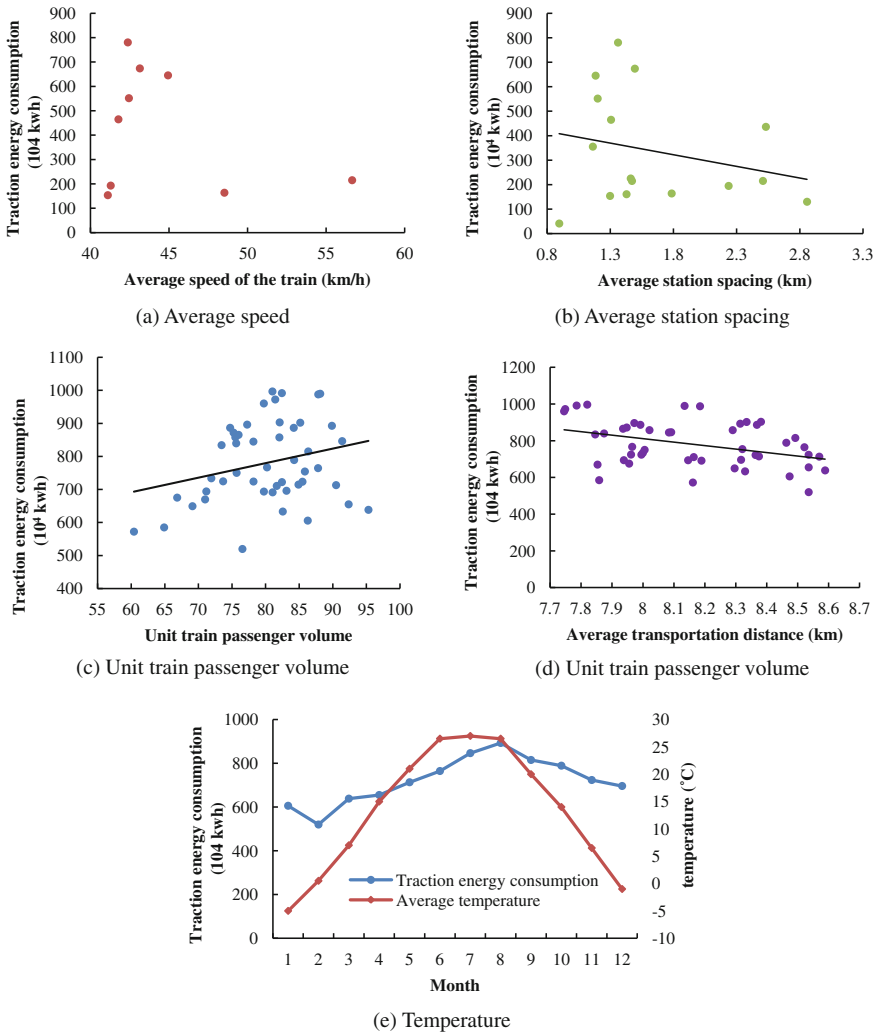
(e) Temperature

Fig. 1 The correlation between several variables with traction energy consumption

$$e = T \cdot (A \cdot V + B \cdot L + X \cdot P + C) \tag{7}$$

where e is the traction energy consumption per t km, kw h/t km; T is temperature variable, as the overall energy consumption of the environmental correction, °C. V and A are set of train running state variables and their parameters set. L and B are respectively the track line attribute variable set and its parameter set; P and X are respectively the track line operation index variable set and its parameter set; C is the constant term.

Data

Based on the urban rail transit statistics database of Beijing, the monthly statistics data of Beijing rail transit, including line 1, line 2, line 5, line 8, line 10, line 13, line 15, Batong line, Changping line, Fangshan line, and Yizhuang line from January, 2011 to July, 2015 are adopted to establish the traction energy consumption model. Then, the modeling database in this paper is established by involving the corresponding basic information of the above lines (e.g., line mileages, station spacing, et al.) and trains (e.g., weight of the train et al.), and the environmental information (e.g., temperature et al.). It is noted that the data from January, 2011 to December, 2014 is used to estimate the model and the data from January, 2015 to July, 2015 is the test data.

Results and Discussion

To quantify the impact of the factors on the traction energy consumption of urban rail transit, the least square method is applied to estimate the parameters of the models, and several statistical tests with regard to multicollinearity, heteroskedasticity and autocorrelated errors are employed to compare and select models. And then, a comparison between the actual value in real-world and the predicted values computing by the proposed models is included in the discussion. Finally, based on the results of the model validation, an extended model involving the season factor is proposed to improve prediction accuracy.

Parameters Estimation

Based on the data mentioned above, the traction energy consumption factor of urban rail transit models for underground line and overground line are calibrated by multiple linear regression methods. Several important statistical tests are performed to determine the nature of the data, such as multicollinearity test, heteroskedasticity test, and autocorrelated errors test [21].

   Caused by the existence of multicollinearity, the effect of a single explanatory variable is not isolated, as the regression coefficients are uninformative and their confidence intervals very wide. Therefore, the multicollinearity test among the explanatory variables is performed to avoid the biased, inefficient and wrongly signed estimates. Here, the variance inflation factor (VIF) proposed by Chatterjee and Hadi [22] is selected to capture the presence of multicollinearity, and then the highly correlated variables are excluded from the explanatory variables used in the final model. From the estimated results, there is no evidence that multicollinearity affects the parameter estimates.

Another important test is heteroskedasticity test. The classical linear regression model requires the variances of non-constant in the regression function are homoskedastic. However, caused by the data usually involved observations from heterogeneous units, the problem of heteroskedasticity is common in cross-sectional analysis. In conducting the Spearman rank correlation coefficient between residuals and predictions, part of the data can be found characterized by heteroskedasticity. Caused by the presence of heteroscedasticity, the ordinary least square is no longer suitable for regression analysis any more, while the estimated results will no longer be efficient.

Furthermore, an important assumption for classical linear regression model is that the error terms in the function should be uncorrelated. This assumption is likely to be violated in regression with time series data, giving rise to autocorrelation. Several methods, such as Autocorrelation Function and the Partial Autocorrelation Function, can be used to detect the autocorrelation among residuals.

Through the above statistical tests, the final parameters calibration results estimated by least square method are summarized in Table 1. The adjusted R^2 with the values over 0.99 exhibit a good statistical fit of the underground line model and overgound line model. Pertaining to the significance of each coefficient, the absolute t-values of the most variables in excess of 1.96 indicate that the estimated results are validated.

Table 1 The parameters calibration results of the traction energy consumption factor model

Variable	Underground line		Overground line	
	Coefficients	t-values	Coefficients	t-values
q	0.005496	6.760981	0.008054	3.033637
$q \cdot v$	−0.00019	−5.42606	−0.00032	−3.00898
$q \cdot v^2$	1.24E−06	3.525392	3.15E−06	2.944291
L_p	−0.00045	−1.48108	0.002101	6.74233
$q \cdot L_p$	2.97E−05	5.722705	−2.2E−05	−2.59907
$m \cdot v^2$	1.2E−07	17.27242	1.02E−07	2.905072
L_s	0.003339	4.717391	−0.00513	−0.82098
T	0.159267	1.522637	−1.32283	−2.72592
$T \cdot v$	−0.00335	−1.54461	0.027637	2.752735
$T \cdot v^2$	2.32E−05	1.559539	−0.00019	−2.78178
T/v	−2.4887	−1.48892	21.06186	2.702061
Sample size	335		192	
Adjusted R^2	0.9904		0.9901	

Where q is the unit passenger volume, person/train; v is the average speed of the train, km/h; L_p is the average transport distance, km; m is the weight of the train; L_s is the average station spacing, km

Model Validation

For evaluating the performance of the forecasting model, the actual values of the traction energy consumption factor in real-world are compared with regression model estimates. The Mean Absolute Percentage Error (MAPE) is selected as the summary statistic indicator to verify the accuracy of the proposed model in this paper. And the calculated formula of MAPE is shown as follows:

$$\text{MAPE} = \frac{1}{n} \sum_{i=1}^{n} \left| \frac{E_a - E_p}{E_a} \right| \times 100\% \tag{8}$$

where E_a is the actual values of the traction energy consumption factor in real-world, kw h/t km; E_p is the predicted values of the traction energy consumption factor, kw h/t km; n is the number of tested data.

Based on the railway energy consumption data from January to July, 2015 in Beijing, the MAPEs for the underground line model and overground line model are calculated as 6.3 and 6.5% respectively, proving a good-fit between the model estimates and the field measurements. And the comparisons between the actual values with predicted values for underground line and overground line are shown in Fig. 2. It is noted that the variations at some data points are significantly more than the others, i.e., the data points 46–50 in underground line, the data points 4–6 and 18–20. Therefore, the model proposed in this section can be further improved.

Model Improving Based on Seasonal Variation

By analyzing the features of each line, it is easy to be found that the traction energy consumption factor of each line show obvious seasonal and periodic features in a year, which is to say, there exist obvious time series characteristics as shown in Fig. 3. It is obvious that the peak traction energy consumption factor appears in June, July and August. Therefore, the season factor, which is defined as a dummy variable, is involved into the model to improve the prediction accuracy. The extended model is shown as below:

$$e' = e + \alpha S \tag{8}$$

where S is the season factor; α is the adjusted parameter of the season.

The estimated results of the extended model are shown in Table 2. And the most parameters in the extended model are statistically significant at the 95% level. Compare to the extended model and the previous model, the extended model have a higher adjusted R^2, which represents the extended model has a better good-fit for
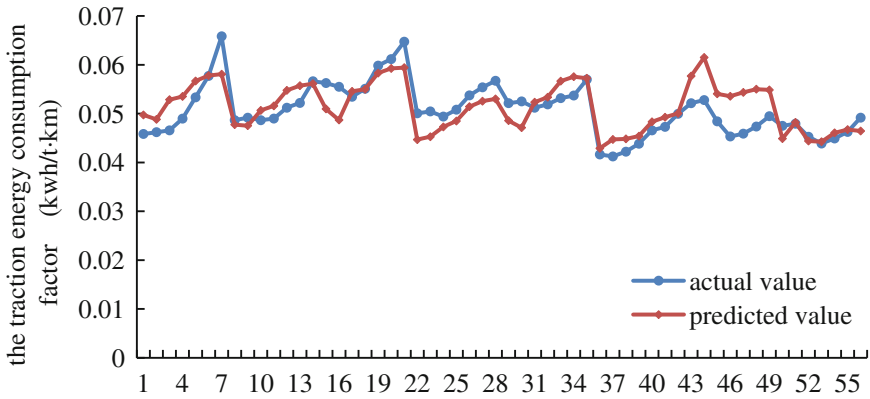
(a) Underground line
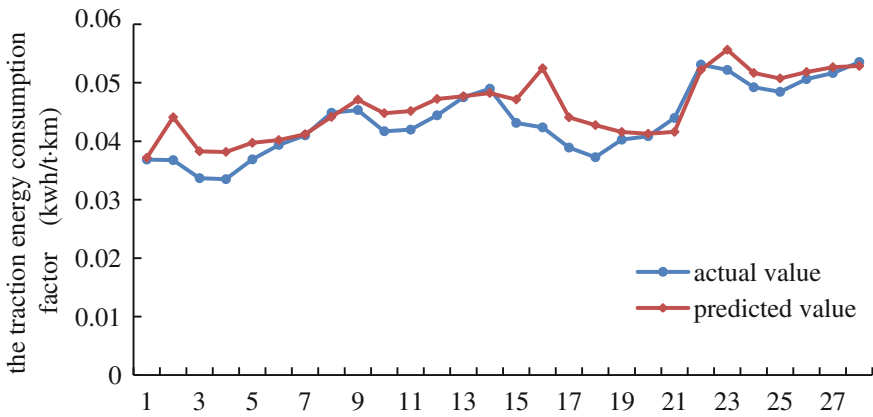
(b) Overground line

Fig. 2 The comparisons between the actual values with predicted values for underground line and overground line

the modeling data. Furthermore, the MAPEs of the above models shown in Table 3 demonstrate that the extended models for underground line and overground line have better prediction accuracy. And the comparisons between the actual values with predicted values computing by the extended model are shown in Fig. 4. It is obvious that the variations of the points (i.e., the data points 46–50 in underground line, the data points 4–6 and 18–20) are effectively reduced. Therefore, the extended model is more effective to evaluate the traction energy consumption factor than the previous model.
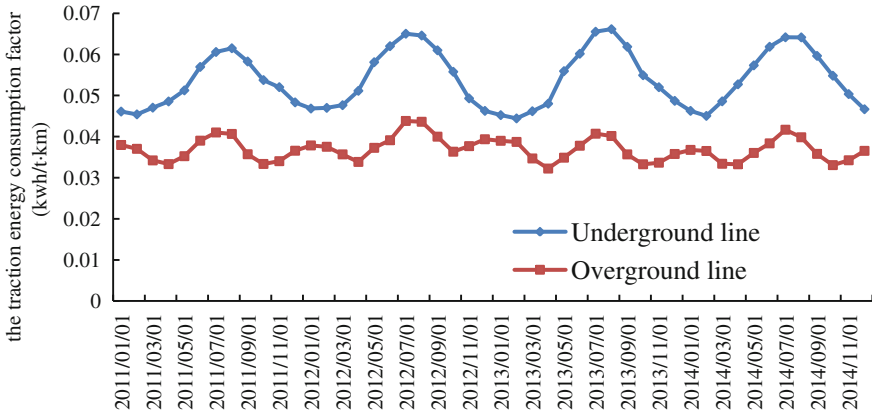
Fig. 3 The comparisons between the actual values with predicted values for underground line and overground line

Table 2 The parameters calibration results of the extended model involving the season factor

Variable	Underground line		Overground line	
	Coefficients	t-values	Coefficients	t-values
q	0.005505	6.955675	0.010276	4.656648
$q \cdot v$	−0.00019	−5.58957	−0.00041	−4.623
$q \cdot v^2$	1.24E−06	3.636197	4.05E−06	4.551738
L_p	−0.00053	−1.78075	0.002366	9.13328
$q \cdot L_p$	3.11E−05	6.137909	−2.1E−05	−3.02744
$m \cdot v^2$	1.22E−07	17.94567	6.92E−08	2.360107
L_s	0.003344	4.852129	−0.0004	−0.07694
T	0.133909	1.312781	−1.21595	−3.03138
$T \cdot v$	−0.00282	−1.33574	0.025446	3.06618
$T \cdot v^2$	1.96E−05	1.350502	−0.00018	−3.1041
T/v	−2.08549	−1.27946	19.32487	2.999337
α	0.00314	4.338413	0.005578	9.226954
Sample size	335		192	
Adjusted R^2	0.9907		0.9914	

Table 3 The parameters calibration results of the traction energy consumption factor model

MAPE	Previous model in Section "Parameters Estimation" (%)	Extended model (%)
Underground line model	6.3	4.7
Overground line model	6.5	5.1

(a) Underground line

(b) Overground line
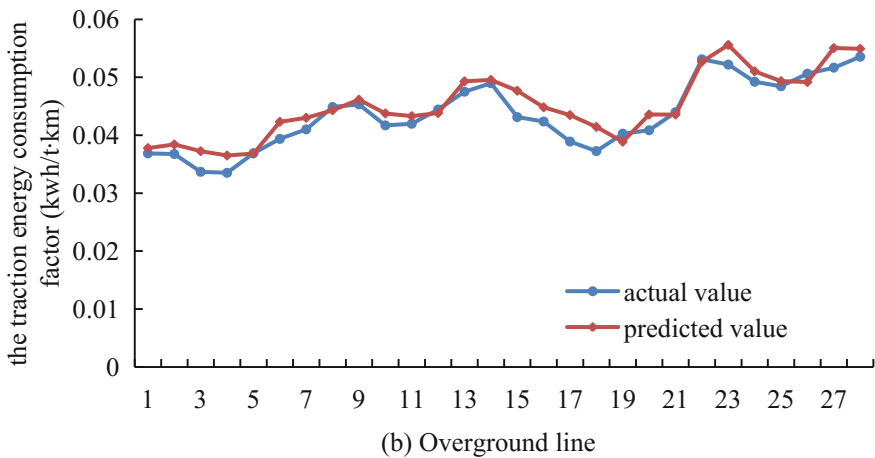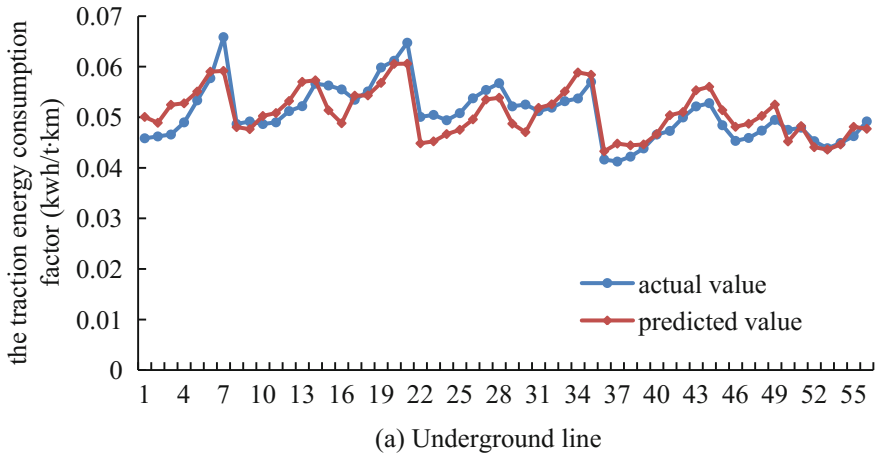
Fig. 4 The comparisons between the actual values with predicted values computing by the extended model

Conclusions

On the basis of dynamic analysis, this article summarized the force status of train during the process of starting, running and braking, and other states, the related factors influencing the energy consumption of train. Finally using statistical regression method to build the unit energy consumption model of rail transit train. Variables of the model are: speed, unit train passenger capacity, the average length of haul, average station spacing, train weight and environmental temperature, etc. Combined with the unit energy consumption showed obvious characteristics of time series, the adjustment of seasonal factors as model variables, established the ground line and the underground line multivariate linear regression model. The energy consumption of

Beijing metro data from January, 2011 to December, 2014 as the foundation, carries on the traction energy consumption statistical regression, and the model was validated by the data from January to July, 2015. The results show that the model to the actual unit ton-km traction energy consumption data fitting degree is high, the interpretation of the variable ability is better, ability to predict better for the future. Therefore, the model can be seen as powerful tools for urban rail transit train traction energy consumption analysis and forecast, helping to optimize operation scheduling, increasing the utilization level of energy consumption.

References

1. Danziger, N.H. 1975. Energy Optimization for rail Public transit system of Washington DC. *Transportation Research Record* 15 (4): 31–39.
2. Hoyt, E.V., and R.R. Levary. 1990. Assessing the effects of several variables on freight train fuel consumption and performance using a train performance simulator. *Transportation Research Part A* 24 (2): 99–112.
3. Hickman, J., D. Hassel, R. Joumard, et al. 1999. Methodology for calculating transport emissions and energy consumption. Cold Starts.
4. Wu, K.Q., Q.X. Sun, X.J. Feng, et al. 2011. Analysis of urban rail transit operational energy consumption. *Journal of Beijing Institute of Technology* 20: 276–281.
5. Peng, Y., M. Lang, and J. Liu. 2010. Reducing energy consumption in China's Railway Freight Transportation. In Wase *International Conference on Information Engineering. IEEE Computer Society, 2010*, 348–351.
6. Liu, H.D., B.H. Mao, Y. Ding, et al. 2007. Train energy-saving scheme with evaluation in urban mass transit systems. *Journal of Transportation Systems Engineering and Information Technology* 7 (5): 68–73.
7. Xu, Xiao-Ming, Ke-Ping Li, and Li-Xing Yang. 2014. Discrete event model-based simulation for train movement on a single-line railway. *Chinese Physics B* 23 (8): 233–239.
8. Liu, P., and Q. Tian. 2012. Analysis of traction energy consumption of urban rail transit. *Shandong Science* 25 (3): 7–11.
9. Fan, H.Q., K.P. Li, and B. Jia. 2014. The influence of network Bottleneck on train energy consumption in railway traffic. *Applied Mechanics and Materials* 631–632: 737–742.
10. Albrecht, T. 2010. Reducing power peaks and energy consumption in rail transit systems by simultaneous train running time control. *WIT Transactions on State-of-the-art in Science and Engineering*, 39.
11. Wang, Y.M. 2011. *Quantification Analysis on the Energy Factors of the Urban Rail Transit System*. Beijing Jiaotong University.
12. Malavasi, G., P. Palleschi, and S. Ricci. 2011. Driving and operation strategies for traction-energy saving in mass rapid transit systems. *Proceedings of the Institution of Mechanical Engineers Part F Journal of Rail & Rapid Transit* 225 (5): 475–482.
13. Bocharnikov, Y.V., A.M. Tobias, C. Roberts, et al. 2007. Optimal driving strategy for traction energy saving on DC suburban railways. *IET Electric Power Applications* 1 (5): 675–682.
14. Gu, Q., T. Tang, F. Cao, et al. 2014. Energy-efficient train operation in urban rail transit using real-time traffic information. *IEEE Transactions on Intelligent Transportation Systems* 15 (3): 1216–1233.

15. Jia, Feng. 2014. *Train Behavior Optimization of Urban Rail Transit System Considering Energy Saving*. Beijing Jiaotong University.
16. Wang, Z.J., F. Chen, and Z.H. Shi. 2013. Energy consumption of Beijing urban rail transit and long-term prediction research. *China Railway Science* 3: 133–136.
17. Xue, Y.B., D.W. Ma, and L. Wang. 2007. Train traction energy consumption calculation method. *China Railway Science* 3: 84–87.
18. Wei, L., L.I. Qun-Zhan, G. Lei, et al. 2010. Study of urban railway energy saving train control optimization based on multi-population genetic algorithm. *Journal of System Simulation* 4: 025.
19. González–Franco, I., and A. García–Álvarez. 2012. Can high-speed trains run faster and reduce energy consumption? *Procedia—Social and Behavioral Sciences* 48: 827–837.
20. Yuan, H.W., and L.Y. Kong. 2012. The factors influencing energy consumption of urban rail transit and measure research. *Urban Rapid Rail Transit* 25 (2): 41–44.
21. Quddus, M.A., N. Harris, and D.J. Graham. 2007. Metro station operating costs: An econometric analysis. *Journal of Public Transportation* 10: 93–107.
22. Chatterjee, S., and A.S. Hadi. 2015. *Regression Analysis by Example*. Wiley.

Calculation Method of Traffic Capacity in Airport Curbside

Yanyan Chen, Nuo Zhang, Haoning Wu, Yao Lu, Liang Zhang and Jun He

Abstract Theoretical analysis and the establishment of airport departure curbside capacity calculation model are described, with the airport curbside as the research object, according to the arrival, stop, and leaving of the vehicle. Taking the Beijing Capital International Airport (BCIA) T3 terminal as an example, this study has evaluated the reliability and accuracy of the model by VISSIM, combined with its traffic characteristics. Departure curbside capacity estimation model derived provides identifiable ground for planning and design of airport departure curbside.

Keywords Airport departure curbside · Traffic flow analysis · Capacity · Traffic planning

Introduction

As the exchange between the countries in the world is becoming increasingly frequent, and regional integration trends are becoming evident, aircraft as the most efficient transport accounts for growing in the proportion of long-distance travel. Therefore, the number of large international airport will increase quickly, which will lead to larger volume of traffic in airport curbside.

Airport is divided into two parts: airside and landside, where landside includes terminals, parking lots, subway stations, and curbsides. In the design process of airports, planning and design of airside and terminal have been a great deal of attention, and formed a clear specification, while the study of landside traffic

Y. Chen (✉) · N. Zhang · Y. Lu
Beijing Key Laboratory of Traffic Engineering, Beijing University of Technology, Chaoyang District, Beijing 100024, China
e-mail: cdyan@bjut.edu.cn

H. Wu · L. Zhang
China Airport Construction Group Corporation, Beijing, China

J. He
Wuhan Communications E&T Information Center, Wuhan, China

Table 1 Annual throughput of BICA (Union: million people)		2010	2011	2012	2013	2014
	Throughput	73.77	78.67	81.93	83.70	86.13

infrastructure is not a lot. At present, the main transportation of domestic large airport landside is taxis, cars, buses and subways. Congestion problems on curbside cannot be ignored with the growing number of passengers. Annual throughput of BICA is listed in Table 1.

The main function of airport curbside is to get off for passengers, so it is non-continuous flow of traffic, and different from urban roads and highways. Thus, capacity of the conventional method is not applicable to airport curbsides. Through the research on the capacity factors of airport curbside, the airport curbside capacity model is established.

Traffic Characteristics in Airport Departure Curbside

Curbside Layout

Different from the urban roads, airport departure curbside is the main channel for social vehicles and the buses getting off and passing. Curbside includes parking lane, weaving lane, and transit lane. Besides, parking lane is used to temporarily park and get off; weaving lane is used for vehicles in and out of the parking lane; and transit lane is used to pass the vehicle.

In general, departure curbside can be decomposed into two or three lane groups, while different types of vehicles occupy each lane group. Each group consists of two or three lanes, a parking lane and a transit lane or a parking lane, a weaving lane, and a transit lane, as shown in Fig. 1.
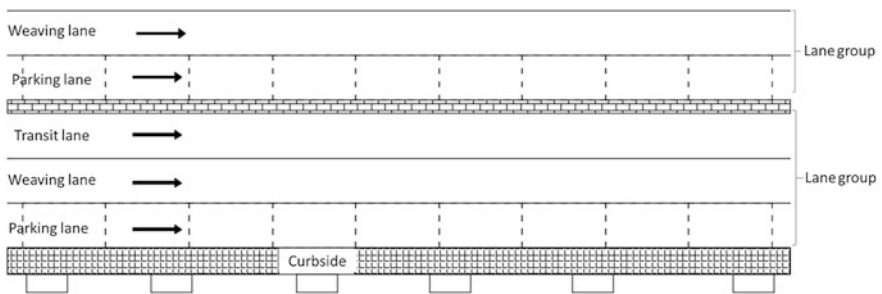
Fig. 1 Schematic diagram of airport departure curbside

Influencing Factors

Parking lane has large influence on the adjacent lane capacity, but models are different in parking time, which leads to influence the adjacent lane at different degrees. Therefore, this article selects models, the parking time of vehicles, and headway as the influencing factors of curbside capacity.

Arrival Rate

Different lane groups for different models of vehicle parking, lane groups are generally divided into the following basic form: the first is for a single model, such as bus, cars and so on; the other is a hybrid dock for several models, for example, the hybrid parking of taxis and buses or social vehicles and taxis.

When parking lane is in the service of a single model, the calculation of the lane group capacity is simple; when parking lane is in the service of two or more vehicle models, it is necessary to get the proportion of each model.

$$R_i = V_i/(V_1 + V_2 + V_3)$$

R_i the proportion of each model on departure curbside (i = 1, 2, 3 representing bus, taxi, and social vehicles);
V_i the volume of vehicle to reach curbside per day.

The arrival of vehicles on departure curbside is mainly affected by plane waves and the travel mode that passenger prefers.

Parking Time

The time of drop-off parking spaces occupied by the vehicle, T_i, can be considered as the summation of the time to drive into the parking spaces, T_{sr}, drop-off time, T_{lk}, and the waiting time to join into outside lanes, T_{se}:

$$T_i = T_{sr} + T_{lk} + T_{sc}$$

Parking time is mainly affected by the carrying capacity of the vehicle, both passengers and bags.

Headway

Headway, t_0, means time interval that two adjacent cars passing around a certain road section in the same direction of traffic [1, 2], and it is an important indicator of the traffic flow. Headway is mainly impacted by the speed and volume of overall traffic.

Number of Parking Spaces

Parking number, N, refers to the number of the standard car parking spaces, which is limited by the length of curbside.

Departure Curbside Capacity Analysis

Parking Lane

Departure parking lane is, in theory, not allowed for passengers to get off, and so the subsequent analysis is based on the following assumptions.

(1) All vehicles are parked in the specified parking position;
(2) Buses account for n_1 parking spaces, other vehicles account for one parking space.

The maximum number of vehicles parking in a parking space per unit time is as follows:

$$c = \sum_{i=1}^{3} R_i \frac{1}{n_i} \frac{3600}{T_i}$$

T_i the time that vehicle occupies the parking space (s) ($i = 1, 2, 3$ representing bus, taxi, and social vehicles);
n_i the number of vehicle that parking spaces occupied;
R_i model's ratio of departure curbside.

T_i, a random variable, can be replaced by the expected value, according to the theory of probability and statistics. So, the service capability, C_t, is as follows:

$$C_t = \sum_{i=1}^{3} R_i \frac{1}{n_i} \frac{3600}{E(T_i)}$$

Transit Lane

Transit lane is similar to city roads, and service capacity of transit lane is as follows:

$$C_g = \frac{3600}{t_0}$$

C_g actual service capacity transit lane (veh/h);
t_0 the minimum headway of transit lane (s).

Weaving Lane

Assumptions

Based on the following four assumptions, the article establishes a method to estimate the service capacity of departure curbside lane:

(1) All vehicles are parked to get off in departure curbside;
(2) There are no stops and driving with illegal phenomenon;
(3) When the parking lane is unavailable, vehicles have to wait on the weaving lane;
(4) Traffic of parking lane is normal, without disturbing by pedestrians, other social vehicles.

Model Construction

With the increase of vehicle in curbside, the vehicle will hinder other vehicles in adjacent lane, at the time to enter and exit the parking Road. Due to this, it will cause the other vehicle speed decreases, increase from the headway and be serious cases congestion, affecting the normal running of the vehicle [3]. Thus, according to the degree of utilization of traffic and the volume into lane group per union time, service capacity of weaving lane is divided into two parts, namely under-saturated traffic of parking lane and overflow of parking lane. Overflow of parking lane means that there is no space to park, so the vehicles later will wait in weaving lane, with a greater impact on the capacity of other lanes; non-overflow of parking lane means to keep the possible capacity.

Parking lane is used for vehicles temporary parking and drop-off, therefore it is faster turnaround. Its turnover is as follows:

$$Q = \frac{3600N}{T}$$

N the number of berths on parking lane (in car count);
T   the average time that vehicle occupies the parking space (s),

$$T = \sum_{i=1}^{3} (R_i \cdot T_i)$$

Non-overflow of Parking Lane

When there is minimum number of vehicles on departure curbside, parking lane meets the needs of the vehicles, so the vehicles can park without wait, that is $Q \geq \lambda$, the service capacity of weaving lane, C_{x1}, is as follows:

$$C_{x1} = \frac{3600}{t_x}$$

λ arrival rate of the lane group;
t_x the minimum headway of weaving lane.

Overflow of Parking Lane

With the increasing number of vehicles on departure curbside, parking lane cannot meet the needs of the vehicles, so the vehicles have to wait, that is $Q < \lambda$, the service capacity of weaving lane, C_{x2}, is as follows:

$$C_{x2} = \left[\frac{3600}{t_x} \right] \cdot f$$

f, the correction factor that parking vehicles impact on the service capacity of weaving lane, $f = \frac{3600 - (\lambda - Q) \times T_w}{3600}$, T_w, the delay time that vehicles wait in the weaving lane(s).

Departure Curbside Service Capacity Model

The service capacity of airport departure curbside is determined by lane's service capacity with different functions. Under different circumstances, they are subject to different impacts of each lane.

Lane Group with Two Lanes

Two-lane lane group contains a parking lane and a weaving lane. When there is a small amount of vehicles on departure curbside, non-overflow on parking lane (that is $Q \geq \lambda$), the service capacities of parking lane and weaving lane determine that of the lane group; when there are so many vehicles that it is overflow on parking lane, this is $Q < \lambda$, the service capacity of parking lane has a significant impact on that of lane group.

$$C = \begin{cases} C_t + C_{x1}...Q \geq \lambda \\ C_t + C_{x2}...Q < \lambda \end{cases}$$

Lane Group with Three Lanes

Three-lane lane group contains a parking lane, a weaving lane, and a transit lane. When there is a small amount of vehicles on departure curbside, non-overflow on parking lane (that is $Q \geq \lambda$), the service capacities of those three lanes determine that of the lane group; when there are so many vehicles that it is overflow on parking lane, this is $Q < \lambda$, the capacity of parking lane has a large impact on that of lane group.

$$C = \begin{cases} C_t + C_{x1} + C_g...Q \geq \lambda \\ C_t + C_{x2} + C_g...Q < \lambda \end{cases}$$

Case Study

This study takes BCIA T3 terminal as an example and investigates the case of the infrastructure and operation of the traffic in a sunny day. Infrastructure includes the length of curbside, structure of lane group, lane width, and the number of parking spaces. Vehicle operating conditions include volume of lane group, speed of vehicles, headway, occupancy, and utilization of parking lane.

BCIA T3 Terminal Curbside Layout

The departure curbside in T3 terminal comprises three lane groups: The innermost lane group near terminal is exclusively for buses and taxis to stop and the middle lane group and the outermost lane group is for private vehicles to stop.

The length of curbside is 400 m, which is available for 80 cars to stop at the same time theoretically. But due to parking spaces unplanned, the utilization rate of parking lane is low (Fig. 2).
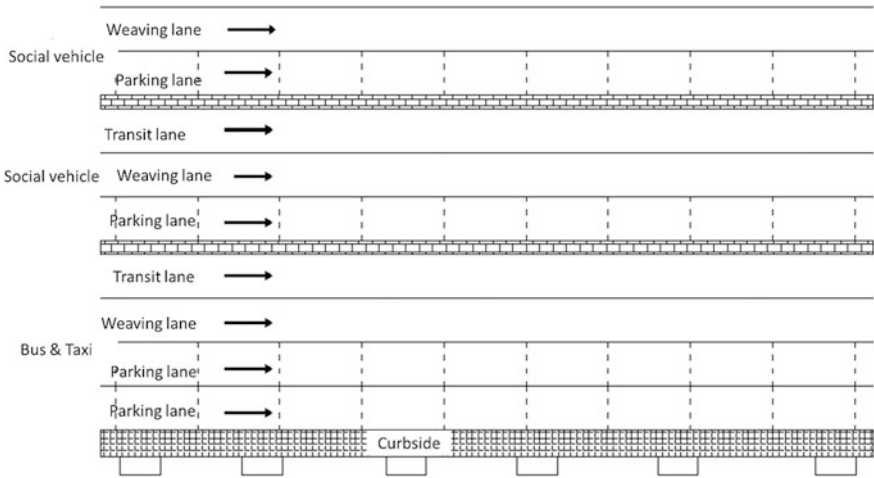
Fig. 2 Schematic diagram of BCIA departure curbside
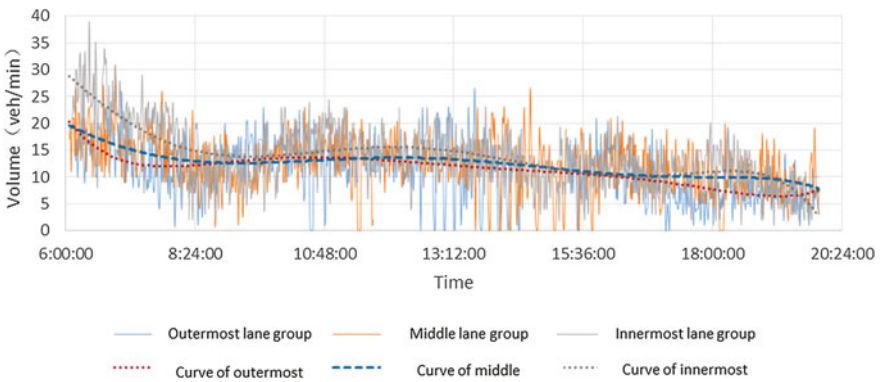
Fig. 3 Arrival of vehicles at T3 departure curbside

Traffic Characteristics

Arrival of the Curbside

Based on the survey data, the arrival of each lane group is derived, and a curve is used to fit the situation of vehicle arrival, as shown in Fig. 3.

As can be seen from the chart, the innermost lane, as lane group for buses and taxis, the arrival of vehicles varies with time. 6: 00, 12: 00 and 19:00 is the peaks of the vehicle arrival. The middle lane and the outermost lane, as lane groups for society vehicles, more stable pattern can be seen.

Table 2 Vehicle arrival of each lane group at T3

	Innermost group lane	Middle group lane	Outermost group lane
Maximum (veh/15 min)	400	292	262
Average (veh/15 min)	204	186	168

Table 3 Proportion and amount of vehicles arriving at T3 innermost lane group

	Taxi	Bus
The amount of arrival vehicle (veh)	10,335	950
The proportion of arrival vehicle (%)	91.58	8.42
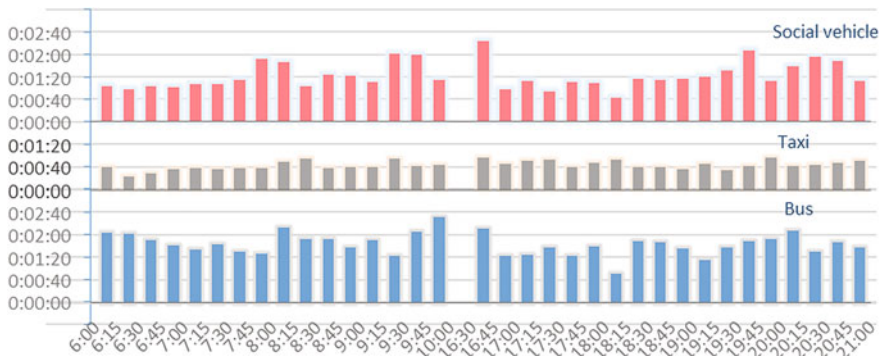
Fig. 4 Parking time of T3 departure curbside

Table 4 Average parking time of vehicles at T3

	Bus	Taxi	Social vehicle
Average dwell time (s)	120	50	85

The arrival of vehicles in each lane group every 15 min is shown in Table 2.

Innermost lane group is specifically for buses and taxis, mixed with a small amount of social vehicles, which is classified as taxis. The proportion of the various models is listed in Table 3.

Parking Time

There are parking time statistics every 15 min, then get the average dwell time of three models in different time periods, as shown in Fig. 4.

From the above chart, we can see that the dwell time of buses and social vehicles is significantly longer than that of taxis. The average dwell time of various models is listed in Table 4.

Table 5 Average headway of each time period at T3 (union: s)

	Flat hour		Peak hour	
	Weaving lane	Transit lane	Weaving lane	Transit lane
Innermost lane group	3.5	3	6	5
Middle lane group	4	3	8	5
Outermost lane group	3		4	

Headway

According to the Traffic statistics of BAIC, peak hours is around 6: 00–7: 00, and flat peak hours 8: 00–9: 00. The result of headway is shown in Table 5.

Curbside Service Capacity Analysis

Through the analysis of data in BCIA T3 terminal, the corresponding parameters above are obtained. Substituting each parameter into the formula model, the results are as shown in Table 6.

Simulation

In order to verify the reliability of the model, curbside capacity model was analyzed, using simulation software VISSIM.

Table 6 Calculation of results of each lane group

	Innermost group lane	Middle group lane	Outermost group lane
Arrival rate (veh/h)	1600	1168	1032
Parking lane turnover (veh/h/lot)	37.1	28.8	22.5
The number of stop (veh/h)	1298.5	1152	900
Correction factor x	0.07	0.93	0.27
Service capacity (veh/h)	1021.3	1231.1	262.5
Capacity of non-overflow of parking lane (veh/h)	2302.8	2128.8	1822.5
Capacity of overflow of parking lane (veh/h)	1394.2	1263.1	922.5

Table 7 Simulation results of airport departure curbside (union: veh/h)

Arrival rate	Capacity simulation		
	Innermost lane group	Middle lane group	Outermost lane group
800	794	775	760
900	873	858	850
1000	970	974	867
1100	1053	1030	867
1200	1158	1121	832
1300	1232	1228	832
1400	1244	1254	841
1500	1202	1240	837
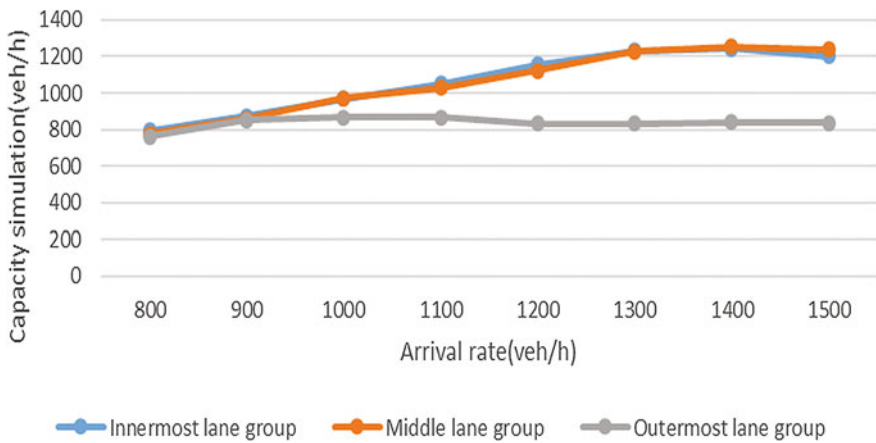
Fig. 5 Simulation results of curbside

(1) Basic data

Basic parameters: dwell time, speed, vehicle type, and proportion.
Input parameters: volume of each lane groups.

(2) Analysis of simulation results (Table 7, Fig. 5)

With the increasing number of vehicles on departure curbside, the simulation capacity of each lane group also is increased. But when the arrival rate reaches a certain value, the capacity of each lane group remains unchanged. The arrival rate reaches 1300 veh/h, the innermost lane group is over-saturated, and capacity remains at 1232 veh/h; when the arrival rate reaches 1300 veh/h, the middle lane group becomes over-saturated and capacity remained at 1228 veh/h; when the arrival rate reaches 900 veh/h, the outermost lane group becomes over-saturated and capacity remained at 850 veh/h (Table 8).

Table 8 Comparison with simulation and theoretical capacity

	Innermost lane group	Middle lane group	Outermost lane group
Simulation saturation capacity (veh/h)	1232	1228	850
The theoretical capacity (veh/h)	1394.2	1263.1	922.5
Relative error (%)	12	1	6

From the above chart, we can see that the error of innermost lane group is larger than that of other lane groups. Due to its special lane group structure, there are two parallel parking lanes, difference parking behaviors in simulation and actual situation. On the whole, theoretical capacity is greater than simulation model. Relative error is small in the middle group and the outermost lane group, indicating that BCIA's capacity has been basically saturated. It means that is necessary to expand or to build a new airport to share increasing passengers.

Conclusion

This paper summarizes the factors that have impacts on the capacity of the airport curbside, including the arrival rate of vehicles, parking time, headway, and the number of parking spaces. On this basis, improving the traditional method of calculating the capacity, the service capacity model for airport departure curbside was built.

Using simulation software VISSIM, the traffic capacity of different arrival rates is simulated. Compared the simulation capacity with the theoretical capacity calculations, we could find that theoretical capacity and simulation capacity in the middle lane group and the outermost lane group comparisons match, while the innermost lane group has two parking lanes with the larger relative error. But the study also has a certain reference value for a multiparking-lane curbside.

References

1. Changlong, Xue. 2005. *Traffic Operation Property of Interleaving on Freeway, 237–254.* Harbin: Harbin Institute of Technology.
2. Xiaohong, Chen, Haifeng, Xiao. 2001. Microscopic simulation of weaving traffic characteristic. *China Journal of Highway.* 2001, 14(sp): 88–91.
3. Yang, Xiaokuan, Jing Cao, and Jian Gong. 2008. Effect of bus stop on the basic link capacity. *Beijing University of Technology* 01: 65–71.

Analysis of an Intersection Based on Active Priority Strategies

Zihao Liu, Xiucheng Guo, Xiaojian Hu and Zhongjun Wu

Abstract This paper focuses on the effect of active priority strategies (APS) in reducing the delay of the bus in the intersection. The research focuses on a typical intersection of Nanjing in China and VISSIM is used to build the model with data collected from the scene. Delay and travel time between two detectors are used to evaluate the effectiveness of the strategy. Besides, the operation effects of social vehicles and buses in current signal timing and two optimized plans programed in Vehicle Actuated Programming (VAP) module of VISSIM are contrasted. It turns out that APS can reduce the average delay and travel time per person greatly and improve the operation efficiency of vehicles in an intersection.

Keywords APS · VAP · Bus priority · Delay in the intersection · Simulation in VISSIM

Introduction

The rapid development of urbanization and automobile industries has caused the problem of land and road resource shortage and the slow construction process of infrastructure makes it even worse. Traffic congestion has affected the sustainable development of the economy and people's daily lives gradually. In the late stage of urbanization, public transport has aroused extensive attention because of its advantages including large capacity, saving energy, and occupying little road resource. It is finally proved home and abroad that giving priority to public transport will improve traffic congestion greatly and promote the sustainable development of the city.

Z. Liu · Z. Wu
School of Transportation, Southeast University, Nanjing 210096, China
e-mail: 1937259449@qq.com

X. Guo (✉) · X. Hu
Southeast University, Nanjing 210096, China
e-mail: seuguo@163.com

Public transport is an essential component of the urban transportation. However, many problems arise during operation which includes large delay, low running speed, low accuracy of punctuality, and bad service, which has made public transport less attractive. Since most of delay, about 60%, happens at the intersection, transit signal priority (TSP) is necessary at the intersection to reduce the delay and improve the efficiency of public transport [1].

When it comes to bus priority, researches focus on two aspects: the priority in space and priority in time. The former focuses on building road facilities such as the bus lane and bus streets and the latter includes passive priority and active priority strategy (APS). Contrast with passive priority, APS will provide service for public transport according to real-time traffic condition, which has a promising prospect. This paper concentrates on APS.

In the research area of APS, Richardson and Heydecker [2, 3] discussed the effect of TSP on delay per vehicle and traffic capacity of intersection and argued that though TSP will lead to the phase miss and rotation, capacity of the whole intersection will not be affected greatly and can be used widely. Liu [4] illustrated two strategies which include early green and extended green with graphical and analytical methods. Francois Dion [5] optimized the APS algorithm and determined whether the current phase should be ended or not and which of the following phase should be executed, which has made APS based on the rule of priority come true. Washington State Department of Transportation put forward three controlling methods of green extension when studying how to carry out TSP on StateRoute7 [6]. Ma and Yang [7] studied the scope of application of early green, green extension, and phase insertion and built the model of delay of bus and delay of public vehicles separately. Liu, H. analyzed the early green and green extension with graphical method and deduced the formula of delay per vehicle of the road with and without priority under the strategy of early green and green extension [4].

Most researchers try to prove the effectiveness of priority strategy with theoretical analysis; however, this paper try to use VAP module with data collected from the studied intersection in VISSIM to justify the effectiveness of APS. Delay per person and travel time are used to evaluate the social and economic benefit that are brought by APS.

Besides, when writing VAP program, this paper not only combines several kinds of APS flexibly, but also puts forward solutions for the condition that several bus apply for priority at the same time, which can be applied to most of the intersections in cities.

Active Priority Strategy and Parameter Optimizing

APS will be actualized after the detectors have detected the coming of buses. When detectors receive the demand for a priority from the bus, priority will be provided to the bus by changing the signal timing of the current phase. APS includes five strategies: green extension, early green, phase insertion, phase jump, and phase

rotation. Strategies that used widely are the first two methods: green extension and early green whose fundamental variables are minimum green time, maximum green time, and green extension time by unit. Theses parameters will be introduced in the following part.

Minimum Green Time

Minimum green time refers to the minimum green time that a phase should own regardless of whether there is a vehicle wants to go through the intersection or not. Several factors should be taken into account when setting the variable [8].

1. Ensure all the vehicles between the detectors and the stop line to drive across the intersection. At that time, the vehicle arrival rate achieves the minimum value in a cycle. It can be expressed as [9]

$$g_{i\,min} = \frac{y_{i\,min}}{Y_{min}(C_{0\,min} - L)}, \quad i = 1 \sim n \tag{1}$$

where,

$y_{i\,min}$ The minimum flow rate ratio of crucial intersection entrance of phase i, and $y_{i\,min} = q_{i\,min}/S_i$;

$q_{i\,min}$ The minimum arrival rate of crucial intersection entrance of phase i in vehicle/s;

S_i Saturation volume rate of crucial intersection entrance of phase i in vehicle/s;

Y_{min} The sum of flow rate ratio of the intersection and $Y_{min} = \sum_{i=1}^{n} y_{i\,min}$;

$C_{0\,min}$ The optimal cycle when arrival rate of crucial intersection entrance of each phase have achieved the minimum value in second and $C_{0\,min} = 2L/1 - Y_{min}$;

L The loss time of each cycle of the intersection in second.

2. Ensue the time that pedestrians can walk through the street safely, which is

$$g_{p\,min} = \frac{l}{v_p} \tag{2}$$

where,

l The width of the road that pedestrians have to go through in mile;

v_p Walking speed of pedestrians in mile/s.

As a result, the minimum green time is

$$G_{i\min} = \max(g_{i\min}, g_{p\min}) \tag{3}$$

Green Extension Time by Unit

Green extension time by unit refers to the interval that the phase has to extend so that the detected bus can arrive at the stop line and go through the intersection after the minimum green time. If no bus is detected during the last few seconds of one phase, then the phase has to change to red. It can be expressed as

$$\Delta_i = \frac{d_i}{v_i} \tag{4}$$

where,

d_i The distance between the detectors and stop line of crucial intersection entrance of phase i in mile;

v_i Speed of vehicles of crucial intersection entrance of phase i in mile/s.

Maximum Green Time

To avoid the increase in delay of the intersection by extending one phase unlimitedly, the ultimate should be set which is

$$G_{i\max} = \frac{y_{i\max}}{Y_{\max}}(C_{0\max} - L), i = 1 \sim n \tag{5}$$

The meaning of each variable is the same as function 1. With the functions above, the green time of each phase can be calculated and the efficiency of the bus will come to the maximum.

Evaluation Index of a Signal Control Intersection

Transit signal priority will change the speed, average delay, and travel time of bus and social vehicles when they go through the intersection. By analyzing these evaluation index outputted in VISSIM, the effectiveness of active priority strategies can be examined. Evaluation index this paper using contains delay per person and travel time [10].

1. Delay per person

Delay refers to the lost time because of traffic congestion and traffic management, which is an important index to evaluate the condition of traffic congestion, the level of service, and the basis of road reconstruction. The aim of APS is to reduce the travel time of passengers in bus and improve the efficiency of public transport. However, the APS will also increase the delay of vehicles with no priority. As a result, analyzing the change of delay per person is essential to evaluate the rationality and effectiveness of APS.

2. Travel time

Travel time means the time a vehicle spends to drive from one section to another section of a road, which includes the parking time during this process. Travel time is a comprehensive index that can be used to evaluate the fluency of the road and to estimate the delay of traffic. Improving the travel speed is the key to enhance the transit efficiency. In the simulation followed, with the detectors set at the start point and end point of the road, travel time can be exported and can be used to identify the effectiveness of APS.
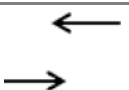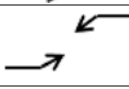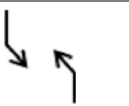
Case Study

The fixed-time control strategy and active priority strategy are analyzed based on the signal intersection of Zhongshan North Road and Shanxi Road of Nanjing and simulation is conducted in VISSIM. The plane and simulation graph is shown as Fig. 1. This intersection is far away from the metro lines and has a complex bus network. The paper tries to implement the APS to this intersection to examine the effectiveness of APS.

Fig. 1 Plane graph of the intersection

Table 1 Volume and signal timing information of the intersection

Phases	Movements	Direction	Car (pcu/h)	Bus (pcu/h)	Arrival rate (pcu/h)	Cycle (s)	Green (s)	Yellow (s)
1	←	WB-straight	312	52	364	160	30	3
	→	EB-straight	364	56	420			
2	↙	WB-left	76	8	84		35	3
	↗	EB-left	124	20	144			
3	↑↓	SB-straight	436	120	556		58	3
		NB-straight	380	124	504			
4	↘↖	SB-left	144	0	144		25	3
		NB-left	144	0	144			

General Situation of the Intersection

The traffic of this intersection is mainly composed of buses and cars. There are specialized lanes for left-turning and right-turning. Since vehicles that turn right are not limited so it is unnecessary to take the right-turning lane into account. According to the survey, the data is summarized in the Table 1.

Optimizing of Current Fixed-Time Signal Control

Since there are some drawbacks in the current signal control plan, Webster signal control strategy can be used to optimize the signal timing. The cycle and green ratio should be calculated with the function 6,

$$C = \frac{1.5L + 5}{1 - Y} \tag{6}$$

In this function, L is the lost time of the cycle which includes the start-up lost time and all-red time. All-red time is 0 and the lost time should be 3 s if no data available. Therefore, for the intersection with four phases, the lost time of a cycle is 12 s. The flow rate ratios of crucial entrance of four phases are 0.177, 0.124, 0.269, and 0.098. As a result,

$$Y = \sum_i y_i = 0.177 + 0.124 + 0.296 + 0.098 = 0.668 < 0.9$$

The flow rate ratio is less than 0.9, and it means that it is reasonable. Furthermore, the cycle can be calculated which is 69.28 s, and it can be assumed as 70 s. Since the green ratio of phase i is proportional to the flow rate ratio y_i, as a result,

$$\lambda_i = \frac{C - L}{C} \times \frac{y_i}{Y} \tag{7}$$

Green ratio of each phase can be calculated and $\lambda_1 = 0.220$, $\lambda_2 = 0.154$, $\lambda_3 = 0.334$, and $\lambda_4 = 0.122$. Green time of each phase is $g_1 = 15$ s, $g_2 = 11$ s, $g_3 = 23$ s, and $g_4 = 9$ s.

The Process of Active Priority Strategy

VISSIM is a simulation software that is created by PTV company in German, which is used to build traffic model and analyze different traffic situations. It is an effective tool to evaluate the effectiveness of traffic planning scheme.

Vehicle Actuated Programming (VAP) module is an additional module of VISSIM for the simulation of programmable, traffic actuated signal controls. Control logic is compiled with simple programming language in a text file. Then, VAP interprets commands of control logic and creates the signal control commands during the simulation for the network [11]. The '*.pua' and '*.VAP' is the key component of VAP module. The '*.pua' file can be programmed in a text file and the '*VAP' file can be compiled in software VisVAP. The function in VisVAP is combined with detectors in the network, which can produce executable file that can be used in VAP module.

The Configuration of Detectors

To execute the APS, induction coil need to be implemented under the pavement near the stop line of an intersection. Normally, the distance between the induction coil and the stop line is about 30–35 m and should be adjusted according to the real situation [12].

Detectors that have implemented are used to detect the coming of bus, and they are 45 m away from the stop line. Forty-five m is decided by the speed of the bus and green extension by the unit, and the detailed computing process can be seen in the calculation of APS parameters part. The position of detectors is shown is the Fig. 2. Detector 1, 2, 6, and 7 are used to detect the coming of the bus from east to west or west to east on straight lanes. Detector 3 and 8 are used to detect the coming of the bus from east to west or west to east on left-turning lanes. Detector 4, 5, 9, and 10 are used to detect the coming of the bus from north to south or south to north on straight lanes. Since there is no bus at left-turning lanes from north to south or south to north, there is no need for detectors in this direction.

Fig. 2 Position of detectors of the intersection

Besides, with the function in VisVAP, bus that has just driven through the detector within 3 s will be detected. In this way, the discrimination of the bus will be improved and more bus will be giving priority to.

The Control Parameters of APS

According to the calculating method in parameter optimizing, the optimized minimum green time, maximum green time, and green extension by unit can be calculated. The optimized cycle is 120 s and the minimum green time and maximum green time of phase 1, 2, 3, and 4 are 22 and 30 s, 20 and 28 s, 20 and 28 s, and 17 and 25 s, separately. The interval of each phase is 3 s.

Normally, the extended green time of a cycle should be controlled in the range of 15–20% of preset fixed-time cycle to ensure that the APS can be finished in a cycle and will not disturb the operation of the subsequent cycle [13]. Assume that the speed of the bus is 40 km/h and the optimized cycle is 120 s, then the extended green time can be set as 24 s which can be distributed into three phases. As a result, the maximum extended time of each phase is 8 s and assume green extension per unit is 4 s, then the distance between detectors and stop line can be calculated and the distance is 45 m.

Programming of APS

1. Logic control program

The control process is shown in Fig. 3. When phase 1 is active, it is necessary to find if the current simulation second is more than the minimum green time. If not,
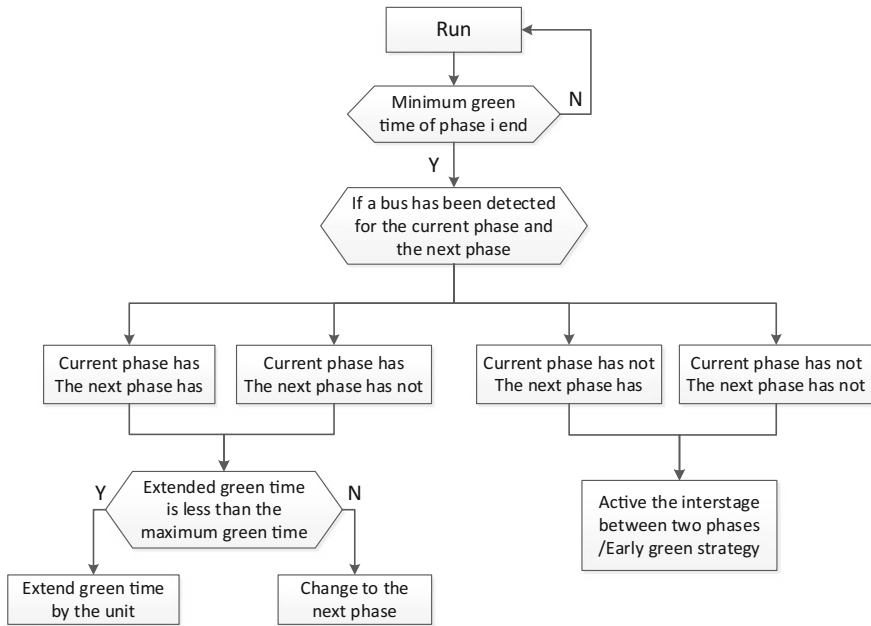
Fig. 3 Control procedure

go on with the simulation second. Else, different measures should be taken according to whether a bus has been detected for the current phase and the next phase. If a bus has been detected in the current phase and the next phase has not, then green extension should be active; otherwise, interstage between two phases should be active which is also the early green strategy for the next phase. Furthermore, if buses are all detected by the current phase and the next phase, which means that there is a conflict between two phases, the green extension will be active for the current phase. And the reason is that according to the relationship between the headway and detected time, only one bus can be detected for each phase and then for the convenience the green extension should be implemented. The processing procedure of other phases is also the same as phase 1.

Besides, by setting the maximum green time, frequent activation of APS will be avoided. The function used in the program is similar to the control logic of the real situation, which ensures the reliability of simulation.

The flow chart above is the key to compile program in VisVAP. The logic control program contains two modules: early green and green extension, which is shown in Fig. 4 [14].

The meanings of all variable are summarized in Table 2.

2. The interstage definition

The definition program of interstage is compiled in a text file with VAP simple programming language. The program is shown in Fig. 5 and it is a '*.pua' file.
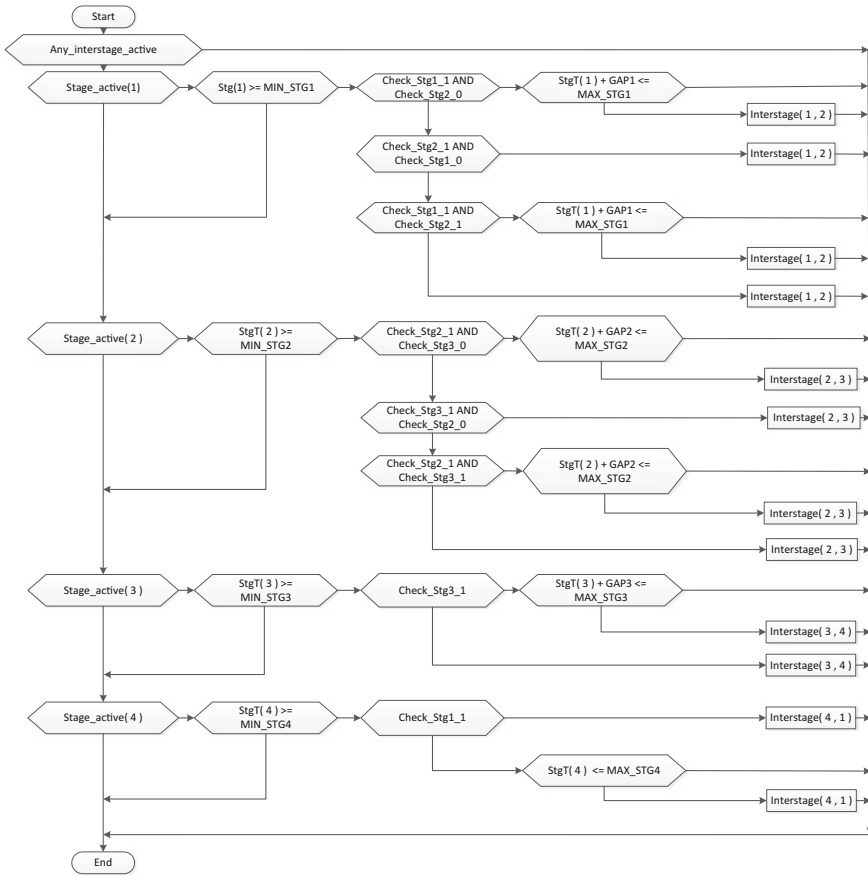
Fig. 4 Control logic of APS in VisVAP

Table 2 Meanings of variables

Symbol	Meaning	Symbol	Meaning
Stage_active (1)	Phase 1 is green	Any_interstage_active	Interstage is active
MIN_STG1	Minimum green time of phase 1	Interstage (1, 2)	Interstage between phase 1 and 2
MAX_STG1	Maximum green time of phase 1	Interstage (2, 3)	Interstage between phase 2 and 3
Check_Stg1_1	Bus is detected in phase 1	Interstage (3, 4)	Interstage between phase 3 and 4
Check_Stg1_0	No bus is detected in phase 1	StgT(4)	Simulation second of phase 4
GAP1	Green extension by unit of phase 1	Headway (1)	Interval since last detection of Phase 1
GAP2	Green extension by unit of phase 2		

```
$SIGNAL_GROUPS
VAP                 VISSIM
$
K1                     1
K2                     2
K3                     3
K4                     4


$STAGES
$
Phase_1             K1
  red               K2    K3    K4
Phase_2             K2
  red               K1    K3    K4
Phase_3             K3
  red               K1    K2    K4
Phase_4             K4
  red               K1    K2    K3

$STARTPHASE
$
Phase_1


$INTERSTAGE
Length [s]        : 3
from Phase        : 1
to Phase          : 2
$
K1                   -127              0
K2                      3            127

$INTERSTAGE
Length [s]        : 3
from Phase        : 2
to Phase          : 3
$
K2                   -127              0
```

Fig. 5 Interstage file

Analysis of Simulation

To examine the effectiveness of optimized fixed-time control strategy and APS, simulations are carried out in VISSIM. The simulation time is 3600 s. The number of vehicles in Table 1 is inputted and travel time, delay per vehicle, and delay per person are used as the observation indexes. The operation results of social vehicles and buses in current fixed-time control strategy and two optimized strategies are summarized in the Table 3 (Figs. 6 and 7).

By comparing the results of fixed-time control strategy and optimized strategy, it can be found that Webster signal control strategy has improved the performance of current fixed-time control strategy greatly. The average delay of the bus and all the vehicles have reduced 52.5 and 46.9% separately and travel time have reduced 27.8 and 31.3% separately compared with the fixed-time control strategy. Besides, assume that the average capacity of social vehicles is 3 persons/veh and capacity of a bus is 30 persons/veh and it turns out that the average delay of optimized fixed-time control strategy has reduced 43%, which reveal the illogicality of the current signal timing.

By comparing the results of optimized fixed-time control strategy and APS, it can be found that the average delay of the bus and all the vehicles have reduced

Table 3 Results of three strategies

Evaluating index	Current fixed-time control	Optimized fixed-time control	APS
Delay of all vehicles(s)	39.696277	18.848937	16.607849
Delay of the bus(s)	39.483221	20.969361	18.259261
Delay per person(s)	217.4335	123.9113	108.9334
Average travel time of all vehicles(s)	66.710914	45.829902	43.580238
Average travel time of the bus(s)	66.522847	48.035659	45.304594
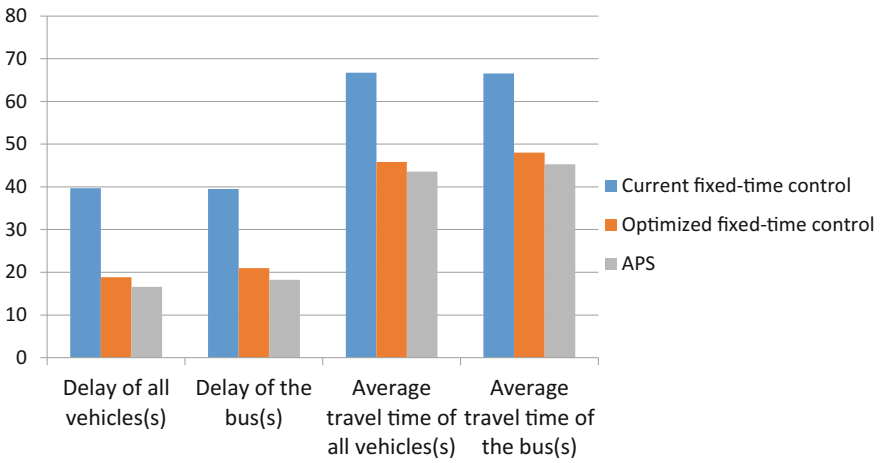
Fig. 6 Delay per vehicle and travel time of three strategies
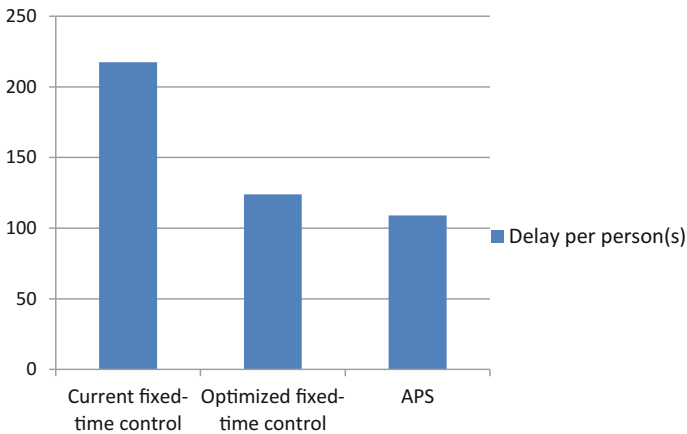
Fig. 7 Delay per person of three strategies

12.9 and 11.9%, separately. Not only the delay of the bus has been reduced, but also the delay of all the vehicles has decreased. The reasons are that for one thing, the optimization of signal timing has improved the traffic capacity of the whole intersection. For another, APS has made positive effect on all the vehicles. Furthermore, the average delay per person has reduced 12.1%, which revealed the effectiveness of APS. In addition, the travel time of the bus and all the vehicles in the intersection have decreased 5.7 and 4.9%, which have proved the effectiveness of APS. Since when APS is active, some social vehicles driving along with the bus also enjoy the same treatment, as a result of which, it is acceptable that the average delay of social vehicles have reduced a little. Overall, by comparing three strategies, it can be easily justified the effectiveness of the active priority strategy.

Conclusion

Transit signal priority strategy is significant to decrease the delay of vehicles and reduce the traffic congestion. The paper focuses on a typical intersection and has realized the active priority strategy with the VAP module in VISSIM. It turns out that the active priority strategy has optimized the performance of Webster signal control strategy. Then, the control strategy the paper puts forward can deal with the situation that buses of the current phase and the next phase apply for priority at the same time. And the APS realized making a sequence for buses who apply for priority at the same time based on the delay per person and improve the traffic capacity a lot. However, there are some shortcomings in this paper that need to be discussed further. On one hand, the strategy the paper put forward is only designed for a single intersection, which will not have a significant impact on traffic in cities and strategies designed for the traffic network is necessary in further work. On the other hand, the strategy that the paper puts forward is suitable for the current traffic volume and will not perform very well if the traffic volume changes a lot. As a result, having a study on the traffic volume that active priority strategy is applicative is also necessary.

References

1. Zlatkovic, M., A. Stevanovic., P. Martin., and I. Tasic. 2012. Evaluation of transit signal priority options for future bus rapid transit line in West Valley City, Utah. *Transportation Research Record: Journal of the Transportation Research Board*, 176–185.
2. Heydecker, B.G. 1983. Capacity at a signal-controlled junction where there is priority for buses. *Transportation Research Part B: Methodological* 17: 341–357.
3. Richardson, A., and K. Ogden. 1979. Evaluation of active bus-priority signals. *Transportation Research Record* 718: 5–12.
4. Liu, H., J. Zhang, and D. Cheng. 2008. Analytical approach to evaluating transit signal priority. *Journal of Transportation Systems Engineering and Information Technology* 8: 48–57.

5. Dion, F., and B. Hellinga. 2002. A rule-based real-time traffic responsive signal control system with transit priority: application to an isolated intersection. *Transportation Research Part B: Methodological* 36: 325–343.
6. Assoeiates, D. 2004. *SR-7 Corridor TSP implementation draft report*. DKS associates: Washington D.C.
7. Ma, W., and X. Yang. 2010. Review of control strategy of transit signal priority. *Urban Traffic*, 70–78.
8. Hu, D., Z. Wu, Q. Fang, and L. Ge. 2009. Research on actuated signal control a single intersection. *Traffic Technology and Economy* 11: 97–99.
9. Zhai, R. 2008. Research on signal timing of transit signal control. *Journal of Chinese People's Public Security University: Natural Science Edition*, 43–46.
10. Ma, H. 2012. *Research on theory and method of bus priority signal coordination control*. China: Chongqing Traffic University.
11. Huang, Q. 2014. *Research on control strategy of bus priority signal at Urban arterials*. China: Southwest Jiaotong University.
12. Zhen, Q. 2012. *Study on the bus priority control technology of Urban road network*. China: Jilin University.
13. Chen, J. 2008. *Study on bus rapid transit priority technologies at signal intersection*. China: Chang'an University.
14. Chen, F., and Z. Wu. 2007. Simulation of induced signal control of an intersection based on VISSIM.

Study on Evaluation Index System of Urban Green Traffic Planning

Jianyou Zhao, Zhongquan Fang and Yang Zhao

Abstract Traffic transformation and upgrading is inevitable request of economy under the new normal. Integrated traffic, safe traffic, smart transportation, and green transportation become the targets of traffic planning in the future; green transportation, as one of the important contents, is attached more importance in urban transportation planning. By comparing the content and requirement of traditional urban traffic planning with urban green traffic planning, putting forward that urban green traffic planning is to improve environmental quality and optimize the use of resources under the premise of meeting the traffic demand and easing traffic congestion, which as the goal layer of urban green traffic planning scheme evaluation. According to the principle of index selection, the evaluation index is determined, and the evaluation index system of urban green traffic planning is established.

Keywords Transportation planning · Green transportation · Planning scheme · Evaluation index

Introduction

In 2014, at the National Conference on Transportation, transportation Minister Yang Chuantang pointed out that in the current period and in the near future to accelerate integrated traffic, safe traffic, smart transportation and green transportation is very demand [9]. Green traffic from academic discussion became the formal national transportation development policy, which has become a major goal of traffic transformation and upgrading in a period of time. Green traffic is a concept, is a city comprehensive transportation system, to reduce traffic congestion, reduce energy consumption and promote environmental friendly, and save the cost of

J. Zhao (✉) · Z. Fang
School of Automobile, Chang'an University, Xi'an 710064, China
e-mail: 1400081104@qq.com; 36126091@qq.com

Y. Zhao
Xi'an 710064, Shanxi, China

construction maintenance [5]. What is more, green traffic is harmonious traffic, including the harmony of traffic and environmental, traffic and resources, and traffic and social. Depending on the content and requirements of green transportation, we need to bring the use of resources and environmental protection into the city traffic planning. The traditional urban transportation planning is to meet the traffic demand and ease the traffic problems. However, green traffic includes more, which is to meet the traffic demand, to improve environment quality, and to optimize the resources utilization [11]. In order to adapt to the basic requirements of green traffic, the evaluation of urban transport planning program also needs to make corresponding adjustments, and evaluation indicators need to reflect the traffic planning in the "green" effect.

Evaluation Objectives

According to the objectives and theory of green traffic planning, the urban transportation planning evaluation index system based on green transportation is divided into three aspects, which include meeting traffic demand, reducing the pollution of the environment, and rational use of resources.

Meeting the traffic demand is the basic goal of the traditional urban transportation planning, also the only goal. It aims to ensure that the planning scheme in implementation can alleviate the contradiction between the growing traffic demand and traffic supply [6]. With the rapid development of urbanization and urban economy, the growth of traffic demand is inevitable. Therefore, to meet the traffic demand is still a basic goal of the urban green traffic planning. The difference is it expands the significance to meet the traffic demand, and traffic supply passively adapts to the change of traffic demand in the traditional traffic planning, which has changed to the coordinated development of traffic supply and traffic demand under the premise of a certain level of service. Under the guidance of this goal, traffic accessibility, travel cost, and fairness are considered as three criteria [10]. Reducing environmental pollution and rational allocation of resources is the new goal, which is not in the traditional urban transportation planning. So it also needs to join the two aspects for the evaluation of the planning program. Problems caused by the traffic exhaust emission, noise, and other environmental pollution problem have become more and more serious. Therefore, in order to achieve the purpose of improving the quality of the environment, on the one hand, policy should encourage people to give priority to the selection of low pollution vehicles to travel; on the other hand, the environmental pollution control should be as the requirement of the whole urban traffic system. What is more, resources use efficiency as a measure of the rational allocation of resources [1].

Evaluation Contents and Procedures

In the process of urban transportation planning, the evaluation of the plan is very important [4]. The traditional urban transportation planning program evaluation focuses on the evaluation of alleviating the situation of traffic congestion and whether to meet the growing traffic demand. The evaluation index [11] mostly is road area per capita, road network density, the proportion of each grade road, etc. The evaluation of urban traffic planning based on green transportation [11] will focus on the view of resource utilization and environmental quality, such as whether the implementation of the scheme reduces energy consumption and reduces pollution emissions in the passenger and freight transport. Only from the comprehensive consideration of traffic function, environmental impact, and resource using, we can determine whether the program is "green" and "green" degree. The evaluation procedure of urban traffic planning based on green traffic is as follows:

(1) To analyze the factors affecting the development of urban green traffic to preliminary determine the evaluation parameters;
(2) To determine evaluation index system, dividing the evaluation system of green traffic for three subsystems from the aspects of the realization of traffic function, environmental impact, and resource consumption, and in each subsystem, to select the typical evaluation index. At the same time, it should be to determine the weight of the index based on the impact degree of these indicators on green traffic;
(3) To determine the planning of the green level through the collection of basic data based on the traffic survey;
(4) By comparing the existing green traffic planning scheme and the green traffic development level, to analysis and propose the improvement solution, in order to achieve the purpose of optimization.

Evaluation Index Selection

The index is the basis of evaluation, which can reflect the "green level" of urban traffic planning. So we should be cautious when selecting indicators. Using qualitative or quantitative methods determines the various indicators, which can show the factors affecting the traffic system. Corresponding to the improvement of urban traffic planning, urban traffic planning scheme evaluation index based on green transportation should consider from the aspects of road network construction, walking traffic, non-vehicle traffic and public transport, parking planning, green transportation mode of good convergence, and logistics park construction.

In order to objectively evaluate the development level and present situation of the urban green transportation, we must give the specific measurable index system and the measure method of the operation. In this paper, on the basis of deep

research of methods to build evaluation system and practical experience, the evaluation index system of green traffic is as follows: firstly, based on the basic connotation of green transportation and the path to realize, according to the international general "target layer-rule layer-index layer" index system logic structure, in reference to successful examples of other cities at home and abroad, combined with the city's own characteristics, proposing alternative indicators set of green traffic evaluation; secondly, according to the principle of index selection and system construction principle, forming basis for evaluation index system; finally, to determine the methods to quantify the index and the evaluation model. The final evaluation index system and calculation method of green traffic are put forward.

According to the three objectives of the green traffic planning, the comprehensive index system of urban green traffic planning scheme evaluation is shown in Table 1.

Table 1 Evaluation index system of urban green traffic planning scheme

Target layer A	Subsystem level B	Criterion layer C	Index level D
Evaluation index system of urban green traffic	Traffic function	Pedestrian traffic	Pedestrian road network density Distance of pedestrian crossing facilities Walking share rate
		Bicycle traffic	Bicycle traffic sharing ratio Bicycle road network density 90% residents travel time
		Public traffic	The coverage rate of public transport stations Average transfer coefficient Proportion of public transport The public traffic network density Proportion of bus lanes Bus standard units
	Environmental pollution	Traffic pollution	The content of PM2.5 in gaseous pollutants The rate of air pollution in sections and intersections The excessive rate of noise pollution in sections and intersections
	Resource use	Land resources use	Road area per capita
		Passenger and freight transport	Fuel consumption per unit of passenger transport The reduction of pollutant emissions per person.km The reduction of pollutant emissions per tons.km
		Trip cost	Total travel cost

Evaluation Indicators of Pedestrian Traffic Planning

(1) Pedestrian road network density (km^2/km)

Pedestrian traffic planning needs to divide the walking unit and determine the function; to different types of walking unit, the walking road network density can reflect the continuity and accessibility of walking road network layout. When the density is insufficient, the road network density can be improved by increasing the corresponding branch, walking special road. The recommended interval values and density values of main walking road network are shown in Table 2.

$$\text{Pedestrian road network density} = \frac{\text{total length of the walking road}}{\text{the total road area of walking unit}}$$

(2) Distance of pedestrian crossing facilities (m)

In order to protect pedestrians and non-motor vehicles crossing the road safely and conveniently, reasonable crossing facilities need to be planned and set up. At the same time, crossing facilities planning and designing should be harmonious with the surrounding environment. The crossing facilities include plane crossing and three-dimensional crossing. Bicycle and pedestrian crossing facilities can be shared, but separate bicycle crossing region should be planned in the case of large bicycle traffic. Tables 3 and 4 are the standard of crossing facilities planning.

Table 2 Walking road network density and walking distance between main walking elements

Walking unit	Network density(km/km^2)	Average interval(m)
Central zone walking unit	≥13	≤150
Public area pedestrian unit	≥11	≤180
Residential walking unit	≥10	≤200
Pedestrian unit of traffic hub	≥10	≤200

Table 3 Maximum distance of each grade road crossing facilities (m)

Road type	Land for resident and social service facilities	Commercial Office	External traffic	Green space and square	Industrial storage
Main spindle road	400	450	500	600	700
The ground road or underground expressway	300	350	400	500	700
Main road	200	200	300	350	600
Secondary roads	150	150	250	300	500

Table 4 Distance of street location to the entrance

Location	Recommended distance	Distance under difficult conditions
The ground bus and rail station entrances or exits	≦80	≦130
The primary and secondary schools, hospital entrance	≦80	≦150
Residential area, large commercial facilities, public activity center entrance	≦100	≦200

Table 5 Table sharing rate level walking

Evaluation level	1	2	3	4	5
Class A and B	≧30	[20,30)	[20,25)	[15,20)	[0,15)
Class C and D	≧35	[30,35)	[25,30)	[20,15)	[0,20)

(3) Walking share rate (%)

"Urban traffic management evaluation index system" (2012) divided cites into ABCD four types based on the urban population, built area, GDP, and other aspects, and it is also clearly pointed out that different types of cities have different walking share rate (as Table 5).

Evaluation Indicators of Public Transport Planning

(1) The coverage rate of public transport stations (%)

The "urban road planning and design standards" clearly pointed out that public transport station service area shall not be less than 50% of the urban area when calculated by radius of 300 m, and shall not be less than 90% of the urban area when calculated by radius of 500 m. It also requests that the bus transfer distance should not be greater than 50 m at the same direction, and the anisotropic transfer distance should not be greater than 100 m.

(2) Average transfer coefficient

The average transfer coefficient of the walker is the sum of numbers of the passenger trips and transfer times divided by the number of transfer passengers. The "urban road planning and design standards" points out that the average transfer coefficient of passengers should not be greater than 1.5 in the "big city" and should not be greater than 1.3 in the small and medium-sized cities.

(3) Proportion of public transport (%)

Public transport is the typical mode of green traffic; we need to calculate proportion of public transport in the evaluation of traffic planning based on green traffic (as shown in Table 6).

(4) The public traffic network density (km^2/km)

Bus network density is defined as the length of the streets in the area unit of urban; the index is an important indicator to measure the degree of urban residents close to the bus network. According to the "urban road traffic planning and design standards," the city's public transport network density is 2 km/2.5 km^2, for the city center or the suburbs the value can be adjusted accordingly.

(5) Proportion of bus lanes (%)

Setting bus lanes is an effective measure to develop public transport priority. In the planning, proportion of bus lanes is shown by the length of the bus lanes divided by the length of the city's total road length (as shown in Table 7).

(6) Bus standard units (million standard units/million)

The index refers to the value of the bus standard units in every million people in city, which reflects the bus service level and traffic structure of urban. For the city with 100 million population [8], the indicators should be 10 million standard units, for the city of 100–300 million population, the index should be 12 million standard units, for city with populations of more than 300 million, the indicators should be 15 million standard units.

Table 6 Bus sharing rate (%)

Evaluation level	1	2	3	4	5
A	≧50	[43,50)	[36,43)	[29,36)	[0,29)
B	≧36	[29,36)	[21,29)	[14,21)	[0,14)
C	≧21	[17,21)	[14,17)	[11,14)	[0,11)
D	≧17	[14,17)	[11,14)	[9,11)	[0,9)

Table 7 Bus lane setting rate (%)

Evaluation level	1	2	3	4	5
A	≧25	[20,25)	[15,20)	[10,15)	[0,10)
B	≧20	[16,20)	[12,16)	[8,12)	[0,8)
C	≧11	[8,11)	[5,8)	[3,5)	[0,3)
D	≧6	[5,6)	[4,5)	[3,4)	[0,3)

Bicycle Traffic Planning Evaluation Index

(1) Bicycle traffic sharing ratio (%)

Share ratio level of bicycle as shown in Table 8.

(2) Bicycle road network density (km/km^2)

Bicycle road network density is a reflection of the bike path length of unit land area (as shown in Table 9).

(3) 90% residents travel time (min)

Approximately 90% residents travel time is that 90% of the residents travel time is less than this value. For the city with different size and different density of the population, the travel time consumption that residents can bear is not the same. In general, the greater the size of the city, people can accept the time of travel will increase [3]. When the usage of city's land becomes more and more reasonable, the large capacity of public transport development is to meet the needs of people with

Table 8 Bike sharing rate (%)

Evaluation level	1	2	3	4	5
A	≥15	[13,15)	[11,13)	[9,11)	[0,9)
B	≥20	[17,20)	[14,17)	[11,14)	[0,11)
C	≥25	[21,25)	[17,21)	[13,17)	[0,13)
D	≥30	[26,30)	[22,26)	[18,22)	[0,18)

Table 9 Recommended values of bicycle road network density and interval

Level	Megacity		Medium and large cities	
	Network density (km/km2)	Average interval(m)	Network density (km/km2)	Average interval(m)
Bicycle main channel	1.5–2.5	800–1300	2.0–3.5	600–1000
Bicycle hub	4.0–8.0	250–500	5.0–8.0	250–400

Table 10 90% residents travel time in different cities

City Scale (10,000 people)	90% residents travel time(min)	Traffic mode
>200	60	Rail transit, rapid transit
100–200	50	Rail transit, bus rapid transit and bus transit
50–100	40	Conventional bus, Bus rapid Transit
<50	30	Conventional bus

different travel purposes, traffic convenience and accessibility will enhance, this value will be reduced accordingly. That is to say, if 90% residents travel time is shortened compared to the previous after the plan implementation, then, the plan is considered more reasonable. 90% residents travel time recommended values are shown in Table 10

Environmental Impact Indicators

From the point of environmental factors, environmental pollution can be divided into four aspects: air pollution, water pollution, noise pollution, and soil pollution [7]. The impact of urban traffic on the environment is mainly reflected in the two aspects of air pollution caused by exhaust emissions and noise pollution in the process of vehicle driving. Therefore, this paragraph focuses on air pollution and noise pollution.

(1) The content of PM2.5 in gaseous pollutants

PM2.5 is recently received a lot of attention to the air pollution index; it refers to particles with a diameter of less than or equal to 2.5 microns in the atmosphere, also known as the lung particulate matter. It is mainly produced by the burning of coal, oil, and other fossil fuels and waste incineration. With the improvement of the level of motorization, people pay more and more attention that individual traffic brings convenience and also brings these problems. Therefore, it is suggested that the content of PM2.5 in road traffic should be included in the evaluation index.

(2) The rate of air pollution in sections and intersections (%)

Because pollutants concentration is the highest on both sides of the road and the intersection, the percentage of the sections mileage and the number of intersections where the air pollution exceeds the limit with the total mileage or the total number of intersections in the city as the quantitative basis of air pollution.

(3) The excessive rate of noise pollution in sections and intersections (%)

According to the "urban regional environmental noise measurement method," the noise during the day shall not be greater than 60 dB, at night not more than 50 db. Select the mileages and the numbers of intersection where noise pollution detection value of expressway, trunk road, trunk road section, and intersection beyond the limit, compared to the total mileage and numbers of intersection number of the noise pollution of road mileage or intersection number city sections of the total mileage or at the intersection, the percentage derived for quantitative data for the estimation of the urban pollution.

Passenger and Freight Transport Evaluation Index

Urban transport in the use of resources, the greatest demand is reflected in the use of fuel and land resources; these are non-renewable resources. Urban traffic is in order to realize the movement of people and objects, therefore subdivided into two aspects of passenger and freight.

(1) Fuel consumption per unit of passenger transport (kg/person.km)

Passenger volume refers to the product of passenger traffic and distance, which as one of the important indicators of passenger traffic. Nowadays, all kinds of traffic tools are mostly using fuel to provide driving force. Therefore, the calculation of the fuel rate not only can provide quantitative guidance from the transportation economics, but also can be used as a measure of the various modes of transport on fuel energy utilization efficiency. Fuel consumption per unit of passenger transport is based on the calculation of passenger volume, whose unit is kg/person/km. In the passenger traffic, we should advocate as far as possible to improve the effective fuel consumption.

$$K_{pt} = \frac{\text{Total fuel consumption of urban passenger transport}}{\text{Urban passenger traffic turnover}}$$

(2) Unit freight volume fuel consumption (kg/km)

Freight turnover refers to the transport distance of freight product with distance and is an important index to describe the freight transportation. Freight traffic is an important part of urban traffic, especially for the city with developed industry and rich resources, the freight traffic is very frequent. In order to save the cost of freight traffic, vehicle is continuously used even if the driver to rest. So freight fuel consumption is very large. The formula for calculating the fuel consumption rate of the unit freight volume is as follows:

$$K_{ft} = \frac{\text{Total fuel consumption of urban freight traffic}}{\text{Urban freight traffic turnover}}$$

(3) Road area per capita (m^2)

In the city, land resources are very valuable, so we should focus on the realization of per unit area of the road to accommodate the passenger traffic volume greater. The formula to calculate road area per capita as follows:

$$\text{Road area per capita} = \frac{\text{Total area of urban road}}{\text{Urban population}}$$

(4) The reduction of pollutant emissions per person.km

Reduction of pollution emissions, especially the reduction of carbon emissions is the target of green traffic, which shows on the reduction of pollutant emissions per person.km in passenger transport.

(5) The reduction of pollutant emissions per tons.km

To achieve reduction in pollution emissions, it is measured by the reduction of pollutant emissions per tons.km.

(6) Total travel cost

The residents travel cost mainly includes the time cost, the cost of the economy, and the influence to the external environment. The time cost is the time that residents selected a certain mode of transport to complete travel consumption, including travel time, transfer time, and waiting time. Economic cost refers to the amount to complete travel, including travel costs, fuel costs, vehicle depreciation costs, walker road use cost (this part can according to the different modes of transport per capita road area with urban planning average land prices), and so on. The cost of influence to external environment refers to the influence on the surrounding air, noise, etc. It is difficult to quantitatively evaluate. Some studies have concluded that the cost of this part can be calculated by taking into account the cost of noise pollution and the cost of air pollution.

$$\text{Residents travel total cost} = \text{time cost} + \text{cost} + \text{external cost}$$

Conclusion

The evaluation index system of urban green traffic planning scheme is constructed, and the traffic function, environmental impact, and resource use are regarded as three target layers. The traffic function is divided into pedestrian traffic, bicycle traffic, and public transport three criteria layer index. Traffic pollution, the use of land resources, passenger and freight transport, and travel cost are regarded as the criterion layer index of environmental impact and resource using. The criterion layer is composed of twenty indicators. The whole index system can reflect the traffic function and the "green" effect of the urban green traffic planning scheme.

References

1. Cheng. Dong-Xiang, Chen Ying et al. 2014. Research on the evaluation model of regional low-carbon transportation. Environmental pollution and control, 36(12).
2. Guo Qin, Wang Yuan-qing 2014. Public Bicycle and Rail Transit Transfer Efficiency Evaluation Index System. Journal of Transportation Engineering and Information, 12(1).

3. Li Peng, Yu Guo-yan 2009. Survey on the Multi-index Comprehensive Evaluation Method. Development & Innovation of Machinery & Electrical Products, 22(4).
4. Lu Hua-pu 2009. Approaches towards Realization of Urban Green Transportation. Urban Transport of China, 7(6).
5. Li, Peng. 2005. *Study on the Evaluated Indices System and the Methods of Implement Decision-Making of the Road Network Plan*. Nanjing: Southeast University.
6. Tang Yun 2015. Study on DPSIR Model of Low Carbon Green Road Transport Development Evaluation. Xi'an: Chang'an University.
7. Tong Gang 2014. Evaluation Index and Evaluation Method on Sustainable Development of Highway Traffic.Transportation Standardization, 42(12).
8. Wu Hui, LI Li 2015. Research on Evaluation Index of Green and Low-carbon Development of Urban Transport. Modern Transportation Technology, 12 (4).
9. Meng-ying, Xu. 2013. *Research for the Improvement of Urban Transportation Planning Method Based on Green Transportation*. Nanjing: Nanjing Forestry University.
10. Yu Ming-yuan, Song Bing 2006. Study on Evaluation Indicator System for Road Transport Planning—Planning Evaluation Indicator System under the Guidance of Scientific Development Concept. Journal of Highway and Transportation Research and Development, 23(8).
11. Zhang Zhi-fang 2015. The Evaluation of the Urban Green Traffic Planning Based on the Value Function Method. Energy Conservation & Environmental Protection in Transportation.

Analyzing the Relationship Between Urban Macroeconomic Development and Transport Infrastructure System Based on Neural Network

Lu Ling, Feng Li and Linhui Cao

Abstract Urban planners have always concerned about the relationship between macroeconomic trend of development and transport infrastructure planning. In the past, urban planners presented qualitative analysis for the construction of transport infrastructure depended on experience. Based on the acquired data of transport infrastructure system and urban macroeconomic, we take the urban development indicators as independent variable and transport infrastructure system indicators as dependent variable. Then, we extract factors and reduce the interference of weak correlation by the method of principal component analysis (PCA). In the aspect of establishing linkage model, BP neural network simplifies topology structures well by adjusting the discrete input, and the minimum error between actual value and predicted value was used as training objectives. Finally, we get a quantitative analysis of the relationship between urban macroeconomic and transport infrastructure. We take Changsha city in China as a case, reducing dimension by PCA and extracting 8 main indicators from original 14 indicators of urban macroeconomic. The precision reaches to 0.99992 and error value is less than 5%. We work out the best weight matrix W_1 and the offset value matrix B_1, B_2 for urban development and transport system in Changsha city. It provides quantitative methodology for policymakers in transport system planning.

Keywords Urban macroeconomic · Transport infrastructure planning · Reduce dimension by PCA · BP neural network model

L. Ling (✉) · F. Li
The Key Laboratory of Road and Traffic Engineering, Ministry of Education
Tongji University, 4800 Cao'an Road, Shanghai 201804, China
e-mail: 13120926990@163.com

F. Li
e-mail: lifeng@tongji.edu.cn

L. Cao
Jiangxi Ganyue Expressway Co., Ltd., 10 Yangju Road, De'an County,
Jiujiang 330400, Jiangxi, China
e-mail: clinhui99@163.com

763

Introduction

The relationship between urban macroeconomic and transport infrastructure system is formed by spatial displacement of individual and objects. A feedback relationship is existed between macroeconomic development of urban and transport infrastructure system [1, 2]. On the one hand, social development required for the root of the traffic demand. The better development of economy, cultural, and political is, the more frequent exchange of position between individuals and objects will be. On the other hand, due to the adjacent effect and cluster effect of social activities, better traffic accessibility will improve the intensity of facilities use, which promotes the strength of the social activities conversely. It is a feedback formation between urban macroeconomic and transport infrastructure system [3].

The relationship between transport infrastructure system and urban macroeconomic in terms of certain influential factors is well researched [4]. The above studies, however, could not cover all the aspects because transportation system is a macroscopic concept. The qualitative analysis can give macroguiding ideology, but specific traffic construction is still unclear. It could mislead the applicability of the transport system and the sustainable development, if only qualitative analysis was adopted. Combining qualitative analysis and quantitative calculation, this paper has put forward transport system planning scheme based on urban development.

It is widely acknowledged that transport infrastructure system planning is directly related to economy, land utilization, population, spatial structure, and some other correlated factors. However, the selection process of these factors is subjective for planner. Various factors have different influence on the transport infrastructure system. Besides, it is easy to overlook some indirect factors which have a great influence on transport infrastructure system [5].

Urban macroeconomic and transport infrastructure have complex relations between each other. However, factors which evaluate them have different correlation. This paper selects out factors by PCA which extracts factors and reduces the interference of weak correlation. Then, we apply the BP neural network to establish linkage model as well as quantitative analysis the relationship between urban macroeconomic and transport infrastructure. It provides quantitative methodology for policymakers in transport system planning.

Data Preparation

Selection of the Independent Variable Indicators

State of urban macroeconomic can be described by a series of indicators including economics, population, land utilization, spatial structure, and the traffic demand associated with transport system. And data is derived from statistical annual.

1. *Economics*

Economics refers to a series of social phenomena and economic attribute [6]. Economics phenomenon has characteristic of economics category and digital embodiment. Main economic indicators include 6 classes: per capita GDP, GDP, industrial added value, agricultural added value, forestry and animal husbandry, and fixed-asset investment and total retail sales of social consumer goods.

2. *Population*

The process of urbanization changes population gathering, and people's behavior derives traffic demand. So, population is an important assess index for urban development. Population indexes are divided into registered population and permanent population which have close relation with urban development.

3. *Land utilization*

Land utilization is the manifestation in the form of urban development, and it is the root of traffic demand [7]. The morphology of land utilization determines product and attraction of urban traffic and traffic mode split. Meanwhile, it rules for city traffic supply and demand model [8]. From a macroperspective, urban area and built-up urban area represent the city scale and the intensity of central city.

4. *Urban spatial structure*

Urban spatial structure refers to the size and shape of city. The spatial distribution is associated with human activities for living, working, entertainment, and relaxation [9, 10]. In the era of integrated traffic information, the cohesion between the various modes of transport infrastructure makes the space of the city integrate. Community office in China is set up according to the population density. In this paper, with the principle of traffic zone division, we use the number of community office to represent the expansion of urban spatial.

5. *Traffic demand*

Traffic demand is the most intuitive performance to transport system. Mutual feedback relationship could be concerned between demand and supply [11]. Traffic problems which relate to the imbalance between supply and demand have become a controversy topic. In this paper, indicators of traffic demand are divided into passenger turnover volumes, freight turnover volumes, and the revolving quantity of society. The independent variable indicators can be shown as follows (Table 1):

Selection of the Dependent Variable Indicators

Transport infrastructure system facilities are close to the spatial distribution of resident trips, which affect the scope and frequency of resident's activities [12]. Traffic mode refers to traffic tools for travels, such as private vehicles (cars), public transport (bus, urban rail, light rail, etc.), and on foot. Dependent variable in this

Table 1 Independent variable indicators

First-grade indexes	Second-grade indexes	Units	Variable name
Economics	Per capita GDP	RMB/person	X_1
	GDP	Billion	X_2
	Agriculture-forestry-animal, industrial add value	Billion	X_3
	Industrial added value	Billion	X_4
	Fixed-asset investment	Billion	X_5
	Total social retail sales	Billion	X_6
Population	Census register population	Ten thousand	X_7
	High school students	Ten thousand	X_8
Land utilization	Urban area	km^2	X_9
	Built-up urban area	km^2	X_{10}
Spatial structure	Community office	a	X_{11}
Traffic demand	Passenger turnover volumes	People · km	X_{12}
	Freight turnover volumes	Billion ton-km	X_{13}
	Revolve quantity of the society	Billion ton-km	X_{14}

paper can be described as 3 types of indicators, including public transport system, motor vehicle demand, and virtual travel intensity.

1. *Public transport system*

Improving the utilization rate of public transport is a means of traffic demand management (TDM). TDM focused on improving the efficiency of the urban transport system through operational improvements which is related to federal policy initiatives in the USA [13]. From the perspective of public transport infrastructure system, we choose bus operation line mileage as an indicator, which refers to all bus travel mileage in one year.

2. *Motor vehicle demand*

The highly motor vehicle ownership attribute to traffic congestion in big city [14], especially private car ownership, which cannot match the limited transport infrastructure. In this paper, we select 2 indicators of "private car ownership" and "private/total vehicle ownership."

3. *Travel intensity of residents*

Travel intensity refers to the residents' average daily travel times which positively correlate to traffic demand. The greater the travel intensity is, the more the resident trips are [15]. From microscopic perspective, travel intensity is the product and attribute between two traffic zones. While from the angle of macroscopic, total actual passenger volume in city can better describe it. "Passenger traffic volume" refers to the total passenger transportation for four kinds of traffic tools including the highway, railway, waterway, and aviation.

Dependent variables are presented in Table 2:

Table 2 Dependent variable indicators

First-grade indexes	Second-grade indexes	Units	Variable name
Public transport system	Bus operation line mileage	km	y_1
Motor vehicle ownership	Private car ownership	veh	y_2
	Private/all vehicle ownership		y_3
Travel intensity of residents	Passenger traffic volume	Ten thousand	y_4

Extraction Factor Based on PCA

PCA is widely used for extracting the main influencing factors of comprehensive variables. In this paper, a large amount of indicators can be used for evaluating urban macroeconomic and transport infrastructure system, while the influence degree of each indicator is not equal. In order to weed out weakly correlated index and improve the accuracy of neural network modeling, this paper applies PCA to reduce the dimension factors and the topology of the simplified model. It lays the foundation of improving the accuracy training of neural network model identification.

Reliability Analysis

This paper poses a reliability test of urban development indicators, taking it as a basis for correction. Usually, the reliability test can be divided into the following grades: 0.60–0.65 is not credible; 0.65–0.70 is the minimum acceptable values; 0.70–0.80 is quite good; and 0.80–0.90 is very good [16].

Cronboch's reliability coefficient is shown as Eq. (1):

$$\alpha = \frac{K}{K-1}\left(1 - \frac{\sum_{i=1}^{K}\sigma_{Y_i}^2}{\sigma_X^2}\right) \tag{1}$$

where

K is the total number of evaluation item;
$\sigma_{Y_i}^2$ is the item variance of Y_i;
σ_x^2 is the sum of the variance for all items.

Factors Extraction Base on PCA

The basic idea of PCA is changing original indexes with strong correlation to unrelated comprehensive indexes, which get more information with fewer factors

[17]. Linear combination is applied to original indexes X_1, X_2, …, X_p, then we get comprehensive indexes as Eq. (2):

$$
\begin{cases}
F_1 = a_{11}X_1 + a_{21}X_2 + \cdots + a_{P1}X_P \\
F_2 = a_{12}X_1 + a_{22}X_2 + \cdots + a_{P2}X_P \\
\cdots \\
F_P = a_{1P}X_1 + a_{2P}X_2 + \cdots + a_{PP}X_P
\end{cases}
\tag{2}
$$

In the Eq. (2), F_i and F_j $(i \neq j,\ i,j = 1,\ldots,p)$ are not relevant, which is named of insufficient information which contained in F_1. We search for the second principal component F_2, the third principal component F_3, … the Pth principal component. a is the correlated coefficient of main factors.

Established Model Based on BP Neural Network

BP (Back Propagation) neural network model is composed of input layer, hidden layer, and output layer, which is a kind of multilayer feed-forward network target for back propagation algorithm. Also, it learns and stores a lot of input–output mapping model [18].

We input learning samples and repeatedly adjust the network weights by back propagation algorithm to make output vector and expect vector nearly as far as possible. When the error sum of squares of output layer is smaller than the specified error, training is finished and network weights are saved. Specific steps are as follows:

(1) Initially, we give the connection weight [w], [b] randomly and threshold θ;
(2) We calculate output of hidden layer and output layer of each unit by the given input–output model;
(3) We calculate the connection weight and threshold;
(4) We calculate error between the predicted values and actual values and judge whether they can meet the specified error. If it is unsatisfied, step 2 is repeated, until the network meets the error precision; otherwise, end the process.

Function Selection

Independent variables and dependent variables are divided into two groups. We take the sum of input values and the weight values as the activation function. During reinforcement learning and adaptive training, we use different activation function model as well as repeated adjust vector and weight. Eventually, we select the tangent S function as activation function for hidden layer and output layer [19], and it can be shown as in Eq. (3):

$$\tan \mathrm{sig}(x) = g(x) = \frac{1}{1 + \exp(-2 * \beta x)} \tag{3}$$

Established the Model

We take dependent variables as units of output and independent variables as units of input and obtain linear summation by nonlinear reaction to comprehensive impression ζ of activation function g [20]. The learning rate is $\eta = 0.05$. Sample pattern is ζ^j. The principle is shown in Fig. 1.

Then, this paper adjusts the result of forecast values and actual values by activation function, until the variance of the predicted values and the actual values is less than 10^{-3}, which meets the specified error precision. Finally, we can get weight matrix of W and bias matrix of B coincide well. It makes sure the value of absolute error less than specified error (5%) between predicted values and actual values, which means weights matrix of W and bias matrix of B can accurately predict evaluation factors of the urban macroeconomic and transport infrastructure system.

Case Study

Changsha is the capital city of Hunan province in China. In recent years, with the development of urbanization and motorization, the scale of Changsha city is expanding and the comprehensive transport infrastructure system is improved correspondingly. Taking Changsha city as example, this paper analyzes the mutual relationship between urban macroeconomic development and transport infrastructure system, which provides methodology for macrourban transport infrastructure system planning.

This paper collects five categories of basic data of urban macroeconomic development as independent variable in the statistical yearbook of Changsha city, including economics, population, land utilization, spatial structure, and traffic

Fig. 1 Neural network topology architecture
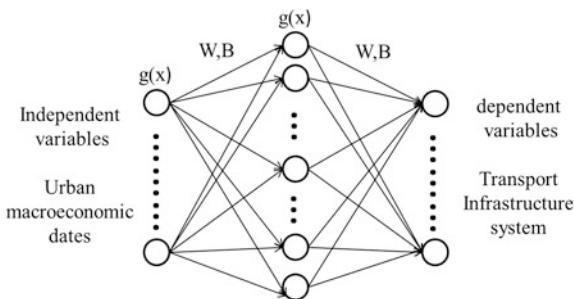
demand. Then, these 5 indicators have been divided into 14 secondary indicators. Meanwhile, we collect three categories of data of transport infrastructure system as dependent variable indicators, including public transport, motor vehicle ownership, and travel intensity, which were further divided into four secondary indicators. We combine PCA with BP neural network and obtain the connection strength weight W and bias values B between the independent variables and dependent variables. All dates are derived from China's urban statistical yearbook of 2014. China's urban statistical yearbook is the comprehensive information yearbook which reflects China's urban economic and social development yearly.

Data Preparation and Analysis

In order to extract the strong correlation data and improve model accuracy strong correlation by PCA, we reduce dimension of 14 kinds of original data of urban development in Changsha city [2] and simplify the neural network topology structure. Finally, we extract 8 main independent variables as X_1, X_2, X_5, X_6, X_7, X_9, X_{12}, and X_{13} as Table 1.

Reliability Analysis

According to the Eq. (1), we get the Cronbach's coefficient of transport infrastructure system data as shown in Table 3, where reliability coefficient is 0.168. After standardizing, it is revised to 0.968, which is better than 0.9. We conclude that internal reliability of each parameter is ideal.

Factor analysis requires strong relationship between variables. When the probability value of each factor is less than the specified significance level, we think it is appropriate for factor analysis. Otherwise, it is considered inappropriate for factors analysis. By KMO and Bartlett's test of sphericity, one can see in Table 4 that the KMO value is 0.709, which is greater than 0.7. Bartlett test value is 568.278, and significance at $P < 0.001$ (less than 0.05) which indicates that there is a high correlation among variables. So the data is considered to be suitable for factors analysis.

Table 3 Reliability statistics

Cronbach's alpha	Cronbach's alpha based on the standard item	
	Alpha	Items
0.168	0.968	14

Table 4 KMO and Bartlett's test

Kaiser-Meyer-Olkin	Bartlett's test of sphericity approx.		
Measure of sampling adequacy	Chi-square	Df	Sig.
0.709	568.278	91	<0.001

Table 5 Explain the total variance

Component	Initial eigenvalue			Extractive sums of squared loadings		
	Total	% of variance	Cumulative %	Total	% of variance	Cumulative %
1	12.439	88.853	88.853	12.439	88.853	88.853
2	1.056	7.543	96.396	1.056	7.543	96.396
3	0.283	2.019	98.414			
4	0.123	0.881	99.295			
5	0.043	0.305	99.6			
6	0.026	0.187	99.787			
7	0.016	0.115	99.902			
8	0.007	0.048	99.951			
9	0.005	0.033	99.983			
10	0.001	0.009	99.992			
11	0.001	0.006	99.999			
12	0	0.001	100			
13	0	0	100			
14	0	0	100			

Factors Extraction Base on PCA

It should meet more than 70% of contribution when extract the comprehensive variables. After spinning, the characteristic value and cumulative contribution rate of each principal component can be shown in Table 5. F_1 is closely tied to per capita GDP, GDP and fixed-asset investment, total retail sales of social consumer goods, passenger turnover volumes, and freight turnover volumes, in which cumulative contribution rate is up to 88.853%. F_2 is composited by census register population and urban area. The cumulative variance contribution ratio of F_1 and F_2 is up to 96.396%. It means the results of factors analysis can be better instead of the original variables. Finally, this paper extracts 8 types of independent indicators from 14 types, by PCA for the next model to cover the most information of the urban development.

Established the Model

The 8 types of independent variables and 4 types of dependent variables are divided into two groups, 85% of sample data is used for training, while the other 15% is used to verify the absolute error values of the model. We take dependent variables y_1, y_2, y_3, y_4 as the four neuron units of output, X_1, X_2, X_5, X_6, X_7, X_9, X_{12}, X_{13} as the 8 neuron units of input and obtain linear summation by nonlinear reaction to comprehensive impression ζ of activation function g. After program debugging, system is divided into 3 layers with 15 hidden-layer neurons.

The principle is as shown in Fig. 2:

This paper adjusts the result of forecast values and actual values by activation function, until the variance of the predicted values and the actual values is less than 10^{-3}, which meets the specified error precision. The results of output variance by BP neural network can be seen in Fig. 3.
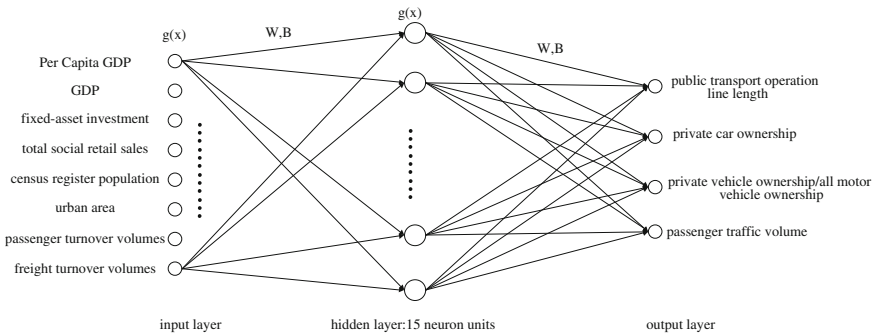
Fig. 2 Neural network topology architecture
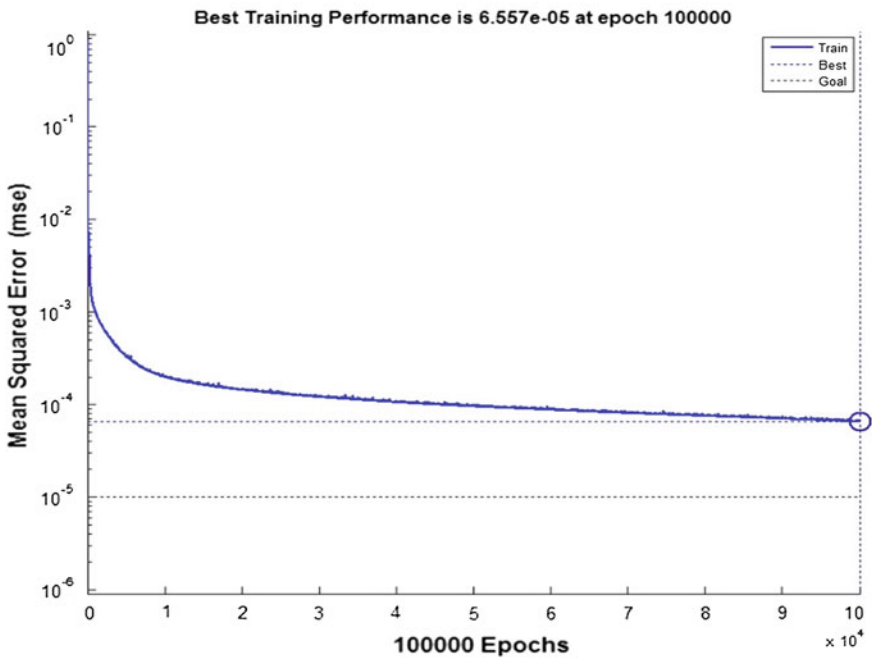
Fig. 3 Output of BP neural network model

This paper gets the value of absolute error between predicted values and actual values of four groups, as shown in the following: $P = \left\{ \begin{array}{cccc} -0.0254 & -0.0223 & -0.0316 & -0.0212 \\ -0.0636 & -0.0568 & -0.0214 & -0.0835 \end{array} \right\}$. The error between predicted values and actual values is less than specified error (5%), which means selected training parameters conformed to the requirements and high prediction precision.

Under the fifteen in the hidden-layer nodes, the fitness between predicted values and actual values is up to 0.99992 which closes to 1. It means weights matrix of W and bias matrix of B coincide well and accurately predict evaluating factors of the urban transport infrastructure system. The weighting matrix of W and offset value matrix of B between predicted values and actual values can be seen as below:

$$
W_1 = \left\{ \begin{array}{ccc}
1.5996 & 2.0447 & -2.2786 \\
-0.4765 & 2.3035 & 2.2275 \\
-2.3943 & -2.3336 & 0.7736 \\
-2.9878 & -1.3520 & -1.9905 \\
2.9455 & -1.9190 & -0.1202 \\
-2.3774 & 2.0226 & 0.9412 \\
2.0225 & -0.4691 & -2.7264 \\
0.1836 & 3.0314 & -1.6493 \\
2.5933 & -1.4167 & -1.7707 \\
-2.4599 & -1.3798 & -1.9117 \\
-0.4096 & -2.7240 & -1.9834 \\
-2.4073 & -2.2635 & -0.5527 \\
2.2879 & 1.7138 & -2.2386 \\
-2.5451 & 0.4969 & 2.1234 \\
1.6880 & 0.0875 & 2.6810
\end{array} \right\}
\quad
B_1 = \left\{ \begin{array}{c}
-3.4486 \\
3.1192 \\
2.5533 \\
1.5458 \\
-1.3203 \\
1.3321 \\
-0.4806 \\
0.0553 \\
0.4186 \\
-1.0613 \\
-1.4527 \\
-2.1408 \\
2.2388 \\
-3.0961 \\
3.6967
\end{array} \right\}
\quad
B_2 = \left\{ \begin{array}{c}
0.3553 \\
-0.4310 \\
-0.3126 \\
-0.3546
\end{array} \right\}
$$

Conclusion Analysis

It can be seen from the bias value matrix B_1 and B_2 that per capita GDP, GDP, fixed-assets investment, and urban area are positively related to bus operation line mileage. On the contrary, total retail sales of social consumer goods, passenger, and freight turnover volumes are negatively related to bus operation line mileage. It shows that investment of public transport cannot keep up with the development of economy and urban spatial structure in the process of urban expansion.

Per capita GDP, GDP, urban area, and total retail sales of social consumer are positive correlated with private vehicle ownership and private/total vehicle ownership, which are negatively correlated to fixed-assets investment. It shows that the growth of economics as well as redistribution of urban structure promotes the number of motor vehicles. However, government leaves no advance for private vehicle in the fixed assets, in the process of the development of Changsha city.

Per capita GDP, GDP, fixed-assets investment, and urban area are positively related to passenger traffic volume, which indicates that the development of macroeconomy accelerates the traffic demand.

Conclusion

Based on the acquired data of transport infrastructure system and urban macroeconomic, taking the urban development indicators as independent variable and transport infrastructure system indicators as dependent variable, this paper finds factors for a stronger correlation by method of PCA. Then, this paper applies the BP neural network to establish linkage model which simplify topology structure by adjusting the discrete input. The minimum error between actual value and predicted value was used as training objectives of BP neural network model. Taken Changsha city as a case, we study the feedback relationship between urban development and transport system based on the 14 independent variables of urban macroeconomic and 4 dependent variable of transport infrastructure. We reduce the independent variable dimension by PCA and extract 8 indexes from 14 indexes which have a greater influence on urban macroeconomic, thus simplifying the input structure and improving the quality of neural network modeling. Finally, we obtain the connection weight matrix W and the bias value matrix B, which meets the specified error precision. In this paper, the BP neural network is adopted to establish dynamic model between the urban macroeconomic and transport infrastructure system. Compared with the traditional methods of qualitative analysis, this method can provide theoretical basis for accurate positioning transport planning and clearly describes the relations between them. Due to the subjectivity of the initial data acquisition by PCA, it brings certain influence to the actual input parameters. Further work should go to the effectiveness of the basic input data.

References

1. Button, K. 2002. City management and urban environmental indicators 40: 217–233.
2. Lei, X., J. Zhang, and J. Li. 2012. A system dynamics model for urban low-carbon transport and simulation in the city of Shanghai, China. *Advances in Information Sciences & Service Sciences* 4 (1).
3. Cascetta, E. 2013. *Transportation systems engineering: theory and methods*. Springer Science & Business Media.
4. Mihyeon Jeon, C., and A. Amekudzi. 2005. Addressing sustainability in transportation systems: Definitions, indicators, and metrics. *Journal of Infrastructure Systems* 11 (1): 31–50.
5. Zhou, S., and X. Yan. 2005. The relationship between urban structure and traffic demand in Guangzhou. *Acta Geographica Sinica* 1: 14.
6. Button, K. 2002. City management and urban environmental indicators. *Ecological Economics* 40 (2): 217–233.

7. Haase, D., N. Schwarz, and Others. 2009. Simulation models on human nature interactions in urban landscapes: A review including spatial economics, system dynamics, cellular automata and agent-based approaches. *Living Reviews in Landscape Research* 3 (2): 1–45.
8. Bartholomew, K. 2007. Land use-transportation scenario planning: promise and reality. *Transportation* 34 (4): 397–412.
9. O'Sullivan, D., and P. Torrens. 2001. Cellular models of urban systems. In *Theory and practical issues on cellular automata*, ed. S. Bandini and T. Worsch, 108–116. London: Springer.
10. Næss, P. 2006. *Urban structure matters*. Routlegde, Abingdon.
11. Richardson, B.C. 2005. Sustainable transport: Analysis frameworks. *Journal of Transport Geography* 13 (1): 29–39.
12. Bunting, T. E., and P. Filion. 1999. Dispersed city form in Canada: A Kitchener CMA case example. *The Canadian Geographer* 43 (3): 268–287.
13. Meyer, M.D. 1999. Demand management as an element of transportation policy: Using carrots and sticks to influence travel behavior. *Transportation Research Part A: Policy and Practice* 33 (7–8): 575–599.
14. Giuliano, G., and J. Dargay. 2006. Car ownership, travel and land use: A comparison of the US and Great Britain. *Transportation Research Part A: Policy and Practice* 40 (2): 106–124.
15. Moudon, A., P. Hess, M. Snyder, et al. 1997. Effects of site design on pedestrian travel in mixed-use, medium-density environments. *Transportation Research Record: Journal of the Transportation Research Board* 1578: 48–55.
16. Cronbach, L. J. 1951. Coefficient alpha and the internal structure of tests. *Psychometrica*. External Resources Pubmed/Medline (NLM) ISI Web of Science CrossRef (DOI) 16: 297–334.
17. Hotelling, H. 1933. Analysis of a complex of statistical variables into principal components. *Journal of Educational Psychology* 24 (6): 417.
18. Rumelhart, D.E., G.E. Hinton, and R.J. Williams. 1988. Learning representations by back-propagating errors. *Cognitive Modeling* 5: 3.
19. Huang, G., Q. Zhu, and C. Siew. 2006. Extreme learning machine: Theory and applications. *Neurocomputing* 70 (1–3): 489–501.
20. Kalogirou, S.A. 2000. Applications of artificial neural-networks for energy systems. *Applied Energy* 67 (1–2): 17–35.

Real-Time Traffic Incident Detection with Classification Methods

Linchao Li, Jian Zhang, Yuan Zheng and Bin Ran

Abstract It is well known that traffic incident detection is essential to intelligent transportation system (ITS) and modern traffic management. Compared to traditional models based on traffic theory, some data mining computational algorithms are believed more appropriate and flexibility for automatic incident detection. In this paper, four classification models were introduced and their parameters were selected by tenfold cross-validation. Using an open dataset their predictive performance was compared based on five criteria. The results show that the classification models perform well to detect traffic incidents and no over-fitting problem. What's more, AdaBoost-Cart and Naïve Bayes models seem to outperform support vector machine and Cart models since they provide superior detection rate. However, they cost long time to train.

L. Li · J. Zhang (✉) · Y. Zheng · B. Ran
Jiangsu Key Laboratory of Urban ITS, School of Transportation,
Southeast University, Nanjing, China
e-mail: jianzhang@seu.edu.cn

L. Li · J. Zhang · Y. Zheng · B. Ran
Jiangsu Province Collaborative Innovation Center of Modern Urban
Traffic Technologies, Nanjing, China

L. Li · J. Zhang · Y. Zheng · B. Ran
Jiangsu Province Collaborative Innovation Center for Technology
and Application of Internet of Things, Nanjing, China

L. Li · J. Zhang · Y. Zheng · B. Ran
Research Center for Internet of Mobility, Southeast University,
Nanjing 210096, China

Introduction

Obviously, non-current traffic congestion and incidents can lead to travel time delays, capacity reduction, and increase of air pollution on roadway network, especially highways which support economic growth. The implementation of intelligent transportation system such as inductive loop detector system is to ensure quick detection of operational problems and to minimize the associated negative effects. Building accurate automatic incident detection (AID) algorithms has been an active research area and experienced a remarkable development in the last few decades. It can be summarized into two types of methodologies: The first type is based on traffic flow theory; the second uses data-mining technique.

Before the emergence of data-mining techniques, some algorithms based on traffic flow theory were widely built to detect incidents. The most popular algorithm was the California algorithm [1] proposed with the idea to compare traffic parameter against a certain threshold value. Also, using statistical theory, some statistical algorithms were developed such as standard normal deviate algorithm [2] and Bayesian algorithm [3]. This kind of AID models estimated traffic characteristics through historical traffic information. Then, the difference between outcome and observed detector data is calculated. Finally, the statistical significance of the difference was evaluated. The representative time series algorithms consisted of moving average model and autoregressive integrated moving average model were also widely used [4].

More recently, data-mining techniques were used successfully throughout the technology and industry. Classification methods as a branch of data mining are to identify which set of categories a new observation belongs. They are suitable to detect traffic incidents and widely used. Typical and popular classification methods include: support vector machine, decision tree, AdaBoost and Naive Bayes [5, 6]. Yuan [7] firstly presented the application of support vector machine in incident detection. Hence, Chen and Wang [8] constructed a support vector machine ensemble, Xiao [9] proposed multiple kennel support vector machine. Ma et al. [10] used support vector machine to detect incidents in cooperation with infrastructure agents. Liu et al. [11] presented the application of Naïve Bayes classifier ensemble to enhance the detection performance.

Up to now, all researchers proved that their AID models had the best performance. However, they lack a comparison of classification methods. It was hard for operators to recognize the best model that they really need. To address these issues, this article compares the four main classification methods to detect traffic incidents with an open dataset.

Data Description

In this study, dataset is from an open database called I-880 which is well known for traffic incident detection. Pretty et al. in University of California, Berkeley, collected the incident data over two periods: One is from February 16 to March 19 and the other is from September 27 to October 29. All data was collected on weekdays during the peak hours (6:30–9:30 a.m. and 3:30–6:30 p.m.) on a section of the I-880 freeway in Hayward, California, which is 9.2 miles long and varied from three to five lanes. There were 35 loop detector stations in experimental segment and data was collected for each 30 s including volume, speed, and occupancy of different lanes.

In this paper, 131,339 cases are extracted from the database and each case contains eight variables: labels, incident time, upstream station speed (km/h), upstream occupancy (%), upstream volume (vehicle per lane per hour), downstream station speed (km/h), downstream occupancy (%), and downstream volume (vehicle per lane per hour). The entire dataset is divided into two parts: a training set and a testing set according to the date. The former contains 88,956 samples including 2438 incident instances and the latter contains 42,383 samples including 1698 incident instances.

Methodology

Support Vector Machine (SVM)

SVM, which was proposed by Cortes and Vapnik [12], became popular in some years ago for solving problems in classification, regression, and novelty detection. The SVM AID can convert the incidents detection problem to find a hyper plane by mapping the origin input into a higher dimensional feature space through a non-linear function ϕ. The output should be two states: 1 for incident scenario and -1 for non-incident scenario. The distance from a point x_n to the hyper plane is given below:

$$\frac{t_n y(x_n)}{\|w\|} = \frac{t_n(w^{\mathrm{T}}\phi(x_n) + b)}{\|w\|} \tag{1}$$

The margin is given by the perpendicular distance to the closest point x_n from the dataset and the goal is to optimize the parameters w and b to maximize this distance. The above algorithm then simply requires to minimize $\|w\|^{-1}$ which is equivalent to solve an optimization problem:

$$\arg\min_{w,b} \frac{1}{2}\|w\|^2$$

$$\text{s.t.} \quad t_n(w^{\mathrm{T}}\phi(x_n)+b) \geq 1, \quad n=1,\ldots,N \tag{2}$$

Sometimes, the incident detection problem is nonlinear separation. We need to maximize the margin while softly penalizing points lie on the wrong side of the margin boundary.

$$\arg\min_{w,b} \left[C\sum_{n=1}^{N}\xi_n + \frac{1}{2}\|w\|^2 \right]$$

$$\text{s.t.} \quad t_n(w^{\mathrm{T}}\phi(x_n)+b) \geq 1-\xi_n, \quad \xi_n > 0, \quad n=1,\ldots,N \tag{3}$$

where the parameter $C > 0$ controls the trade-off between the slack variable penalty and the margin. ξ_n is the relaxation variable that is used to optimize nonlinear separation. According to Lagrange theorem, the above problem can be transformed into its dual optimization problem:

$$\tilde{L}(a) = \sum_{n=1}^{N}a_n - x\frac{1}{2}\sum_{n=1}^{N}\sum_{m=1}^{N}a_na_mt_nt_mk(x_n,x)$$

$$\text{s.t.} \quad 0 \leq a_n \leq C \quad n=1,\ldots,N$$

$$\sum_{n=1}^{N}a_nt_n = 0 \tag{4}$$

The final decision function can be defined as follows:

$$f(x) = \text{sign}\left(\sum_{n=1}^{N}a_nt_nk(x_n,x)+b\right) \tag{5}$$

where $k(x_n,x)$ is kernel function. Linear, polynomial, RBF, and sigmoid are the four types of kernel functions which are widely used. Readers can find more mathematical details on SVM in book as reported by Bishop [13].

Naïve Bayes (NB)

Naïve Bayes methods are a set of supervised learning algorithms based on applying Bayes' theorem with the assumption of independence between every pair of features. The detail procedures are as following:

1. Given a class variable y: 1 for incident scenario, -1 for non-incident scenario and independent variables x_i, $i = 1, \ldots, n$. Bayes' theorem states the following relationship:

$$P(y|x_1, \ldots, x_n) = \frac{P(y)P(x_1, \ldots, x_n|y)}{P(x_1, \ldots, x_n)} \tag{6}$$

2. Based on the independence assumption that the posterior can be represented as:

$$P(y|x_1, \ldots, x_n) = \frac{P(y)P(x_1, \ldots, x_n|y)}{P(x_1, \ldots, x_n)} \propto P(y) \prod_{i=1}^{n} P(x_i|y) \tag{7}$$

3. Predict data cases as class value y that its independent variables maximize the posterior probability as follows:

$$\hat{y} = \arg \max_{y} P(y) \prod_{i=1}^{n} P(x_i|y) \tag{8}$$

Despite its unrealistic independence assumption, NB classifier is surprisingly effective in practice since its classification decision may often be correct even if its probability estimates are inaccurate [14].

Cart Algorithm

Cart, proposed by Breiman et al. [15] in 1984, is considered as a binary decision tree algorithm, that is, beginning with the root mode that contains the whole learning sample and constructed by splitting a node into two child nodes, repeatedly. The development of a Cart model mainly contains two steps: split and prune. The splits are selected using the Gini Index and the obtained tree is pruned by cost-complexity pruning.

Gini index is an impurity-based criterion that measures the divergences between the probability distribution of the target attribute's value. It is defined as:

$$i(t) = \sum_{j \neq i} p(k|t)p(l|t) \tag{9}$$

where $i(t)$ is a measure of impurity of node t and $p(k|t)$ is the node proportions (i.e., the cases in node t belonging to class k).

Consequently, the evaluation criterion for selecting the greatest improvement in the purity score of the resultant nodes is solving the following maximization problem:

$$\arg\max_{x_j \leq x_j^R, j=1,...,M} [i(t) - p_R i(t_R) - p_L i(t_L)] \tag{10}$$

where p_R and p_L are the proportions of observations of the parent node t that goes to the child nodes t_R and t_L, respectively, and j is the number of variables. X is variable matrix with M number of variables x_j and N observations. x_j^R is the best splitting value of variable x_j.

Pruning is a mechanism to create a sequence of simpler trees, through cutting off increasingly important nodes. The primary task is to find the optimal proportion between the tree complexity and the misclassification error. The node misclassification cost can be defined as:

$$r(t) = 1 - p(j|t) \tag{11}$$

The tree misclassification cost can be defined as:

$$R(T) = \sum_{r \in T} r(t)p(t) \tag{12}$$

The cost-complexity measure for subtree T can be defined as:

$$R_\alpha(T) = R(T) + \alpha(\tilde{T}) \rightarrow \min_T \tag{13}$$

where $\tilde{T}$ is total sum of terminal nodes in the tree; $\alpha(\tilde{T})$ is complexity measure that depends on $\tilde{T}$; and α is parameter that measures how much additional accuracy is added to the tree to warrant additional complexity. During the pruning process, the value of α will gradually increased from 0 to 1. For each α, a subtree can be found to minimize $R_\alpha(T)$. Therefore, by gradually increasing α, a sequence of pruned subtrees from the saturated tree can be generated. α can be found through tenfold cross-validation. It is a method randomly partition training set into ten equally size groups. Then nine are used to train the classifier, and the other is used for test. The detail information about ten cross-validations is presented in book of McLachlan et al. Esposito et al. proposed 1-SE rule to simply the tree based on cost-complexity measure. If complexity parameter within one standard error of achieved minimum is marked as the best [16].

AdaBoost-Cart (ACT)

Boosting, a machine learning ensemble meta-algorithm for reducing both bias and variance in supervised learning by converting weak learners into a strong one, was first proposed by Schapire [17] in 1990. The advantage of boosting is it applies the weak classifiers repeatedly on part of training data with adaptive weights and makes

maximum use of each weak classifier. In this paper, the best known AdaBoost algorithm is the most common implementation of boosting, which is introduced to detect traffic incidents. Usage of classification tree as weak learners ACT is built up. The algorithm is as follows:

First step: Start with $w_n^1 = 1/N$ for $n = 1, 2, \ldots, N$. Where n is the number of training observations; w_n^1 is the initial weight.

Second step: Repeat for $m = 1, 2, \ldots, M$:

(a) Fit a classification tree $y_m(x) \in \{-1, 1\}$ using w_n^m on the training data. Where 1 represent for incident and -1 for non-incident.

(b) Compute the quantities: $\varepsilon_m = \dfrac{\sum_{n=1}^{N} w_n^m I(y_m(x_n) \neq t_n)}{\sum_{n=1}^{N} w_n^m}$ and $\alpha_m = \frac{1}{2} \ln\left(\frac{1 - \varepsilon_m}{\varepsilon_m}\right)$. Where $I(y_m(x_n) \neq t_n)$ is an indicator function and equals 1 when $y_m(x_n) \neq t_n$ and 0 otherwise.

(c) Update the data weighting coefficients $w_n^{m+1} = w_n^m \exp\{\alpha_m I(y_m(x_n) \neq t_n)\}$ and normalize them.

Third step: Output the final incidents detection model (Ada trees) $Y_M(x) = \text{sign}\left(\sum_{m=1}^{M} \alpha_m y_m(x)\right)$.

The training set is classified with a boosting ensemble of M trees in order to find the best M cross-validation that estimates how accuracy a model will perform is used. Obviously, the lower error of a tree is, the higher its weight in the final ensemble.

Performance Criteria

In the exist literature, there are four commonly performed indices that are widely used: accuracy (ACC), detection rate (DR), false alarm rate (FAR), and mean time to detect (MTTD). Their definitions are as follows:

$$\text{ACC} = \frac{\text{The number of true reports}}{\text{The total number of cases}} \tag{14}$$

$$\text{DR} = \frac{\text{The number of true incident reports}}{\text{The total number of incidents}} \tag{15}$$

$$\text{FAR} = \frac{\text{The number of false incident reports}}{\text{The total number of cases}} \tag{16}$$

$$\text{MTTD} = \text{The time AID costs to detect incidents in the data set} \tag{17}$$

Receiver operating characteristic (ROC) is a graphical plot that illustrates the performance of a binary classifier system as discrimination threshold is also introduced. ROC curves typically feature true positive rate on the y-axis, and false

positive rate on the *x*-axis which means that the top left corner of the plot is the ideal point—a false positive rate of zero and a true positive rate of one. This is a situation that is not very reality but it does mean that a larger area under the curve (AUC) is usually better.

Result

Parameter Selection

In this section, we take the training set to calculate the best parameters of the four different models with tenfold cross-validation (Fig. 1).

For SVM model, radial kernel function is utilized. In modeling process, grid search method is applied to find the optimal parameters γ and C. A large search area with a smaller step may increase the probability to get the best parameter but cost longer computation time, and vice versa. Therefore, in this experiment, firstly we conduct a large search with a big step to confirm the interval that contains the best parameter. After that a smaller search with a smaller step is used. The result can be found in the Fig. 2. The former uses the large search area and the latter is more precise. Finally, we find SVM with radial kernel function when $C = 53$ and $\gamma = 1.0$ has the highest tenfold cross-validation accuracy shown in Table 1.

For Cart algorithm, the interpretation of a tree with large number of splits and terminal nodes is increasingly difficult. Furthermore, it can cause over-fitting problem when the number of splits levels approaches the number of observation without constrains. In this analysis, the optimal tree structure selected with 1-SE rule has 11 terminal nodes. The result is shown in Table 1 and in Figs. 2 and 3.

For ACT, as mentioned above, it implements AdaBoost with classification and regression trees as base classifiers. It is necessary to calculate the number of trees to use. As shown in Fig. 4, after dashed line (30 trees), the three criteria become steady, so 30 is a proper number for ensembles.
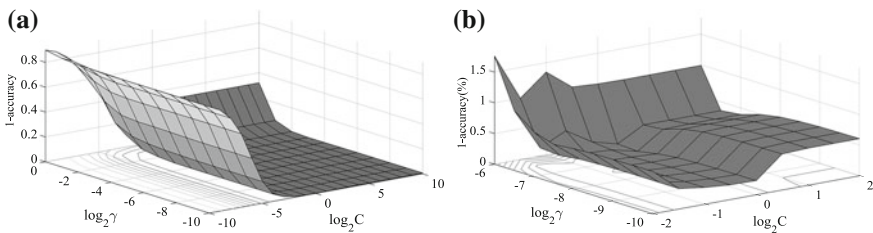
Fig. 1 Result of grid search method
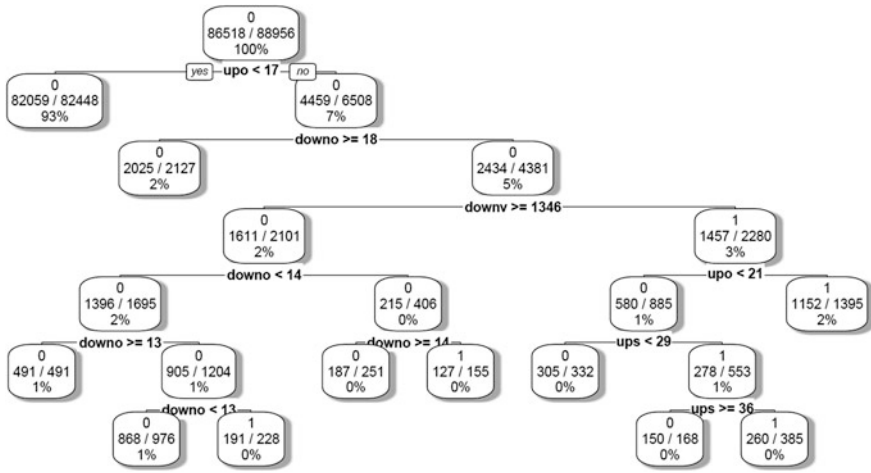
Fig. 2 Result of the Cart with the best parameters

Table 1 Error of different number of splits

Number of splits	0	4	6	**10**	12	15
α	**0.087**	**0.028**	**0.026**	**0.011**	0.010	0.010
Standard error	0.020	0.016	0.015	**0.014**	**0.014**	**0.014**
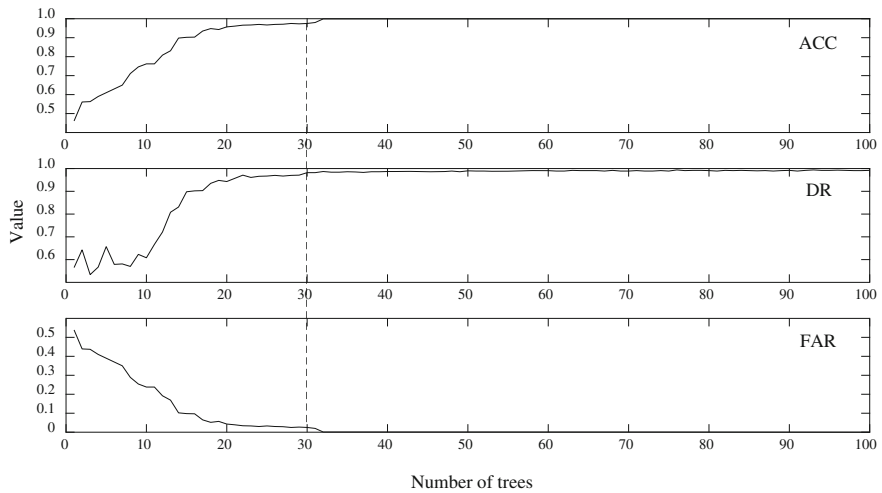
Fig. 3 Criteria of different number of trees in the ACT models

Fig. 4 ROC of four different
AID models
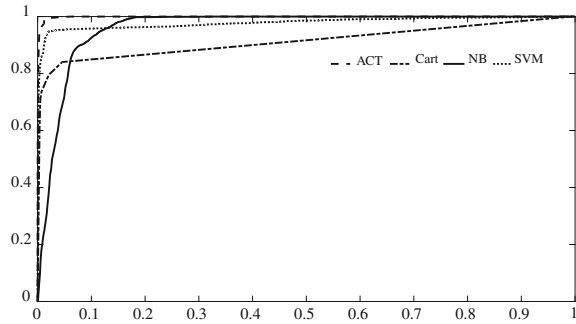
Comparison

The performance measures of training and testing are derived from the average of
the ten confusion matrixes and shown in Table 2. Also the ROC curves for the four
models are shown in Fig. 4 and the area under the ROC (AUC) is calculated that
can be seen in Table 2. ROC curve is a graphic plot that illustrates the performance
of binary classifiers. From Fig. 4, obviously, ACT algorithm is superior to others.

The test result of NB is interesting. It has the most powerful capacity to detect
about 82.7% incidents but it tends to detect some non-incidents as incidents and
give false alarm to the operators seen form high FAR (0.137). What's more, NB
costs the longest test time (7.530) that is a little longer than test time of SVM
(6.740) and far more than the test time of Cart (0.060) and ACT (0.840). It is logical
to discover NB is not over-fitting that is consistent with the perspective of Bishop
who thinks Bayesian approach can avoid over-fitting problem [13]. Nevertheless,
we find the FAR of NB is a little high.

When considering Cart and SVM algorithms, it is true that their ACC is high
enough but their DR is around 50–60% which may not meet the requirement of
operators. It is essential that operators actually pay less attention to this criterion.
They think DR is more significant because the consequence of a true incident
unrecognized is more serious than a non-incident is false alarm.

Table 2 Summary of the training–testing results

Criteria	SVM		NB		Cart		ACT	
	Training	Testing	Training	Testing	Training	Testing	Training	Testing
ACC	0.991	0.961	0.937	0.893	0.987	0.972	0.995	0.984
DR	0.725	0.589	0.838	0.827	0.710	0.597	0.882	0.814
FAR	0.002	0.021	0.060	0.137	0.005	0.012	0.002	0.009
MTTD (s)	190.730	6.740	0.150	7.530	1.030	0.060	12.650	0.840
AUC	0.978		0.962		0.913		0.999	

In Table 2, it shows the ACT model has high DR and ACC but low FAR. ACT possesses the highest ACC that means they have strong ability to correctly recognize the cases in the whole testing set. The DR is 0.814 that proves it can select about 81.4% incidents from all incidents in the testing set which is only lower than NB (0.827). Moreover, the low FAR also indicates it is suitable for incidents detection.

Conclusion

The paper presented a comparison of four classification algorithms for highway traffic incident detection using the real traffic data (e.g., speed, volume, and occupancy) from I-880. The proposed AID methods consist of training phase and testing phase whose detail procedures are introduced in the methodology section, and their key parameters are calculated. After evaluating four AID algorithms, we do a comparison experiment to study their performance and discuss the over-fitting problems with tenfold cross-validation. The results clearly indicate there is no over-fitting problem for data-mining techniques to use as incidents detection AID models. However, there are also some sophisticated AID models that are built based on modified SVM, NB, Cart, and ACT algorithms which we do not know whether they have over-fitting problem or not. Actually, the origin simple classification methods are accurate enough to detect traffic incidents. NB and ACT models perform quite well compared to SVM and Cart models for this type of problem. ACT costs short time in test phase but long time in training time whereas NB is just the reverse.

A possible improvement to consider in the future is the imbalance of the dataset which may affect the accuracy of classification models. It is better to study the method to balance the data.

References

1. Payne, H.J., and S.C. Tignor. 1978. Freeway incident-detection algorithms based on decision trees with states. *Transportation Research Record* 682: 30–37.
2. Dudek, C.L., C.J. Messer, and N.B. Nuckles. 1974. Incident detection on urban freeways. *Transportation Research Record* 495: 12–24.
3. Levin, M., and G.M. Krause. 1978. Incident detection: A Bayesian approach. *Transportation Research Record* 682: 52–58.

4. Ahmed, S.A., and A.R. Cook. 1982. Application of time-series analysis techniques to freeway incident detection. *Transportation Research Record* 841: 19–21.
5. Wu, X., and V. Kumar, eds. 2009. *The top ten algorithms in data mining*. CRC Press.
6. Han, J., M. Kamber, and J. Pei. 2011. *Data mining: Concepts and techniques*. Elsevier.
7. Yuan, F., and R.L. Cheu. 2003. Incident detection using support vector machines. *Transportation Research Part C: Emerging Technologies* 11 (3): 309–328.
8. Chen, S., and W. Wang. 2009. Decision tree learning for freeway automatic incident detection. *Expert Systems with Applications* 36 (2): 4101–4105.
9. Xiao, J.L., and Y. Liu. 2012. Traffic incident detection using multiple-kernel support vector machine. *Transportation Research Record* 2324: 44–52.
10. Ma, Y., M. Chowdhury, M. Jeihani, and R. Fries. 2010. Accelerated incident detection across transportation networks using vehicle kinetics and support vector machine in cooperation with infrastructure agents. *IET Intelligent Transport Systems* 4 (4): 328–337.
11. Liu, Q., J. Lu, K. Zhao, and S. Chen. 2014. Naive Bayes classifier ensemble for traffic incident detection. In *Transportation Research Board 93rd Annual Meeting*, Washington, D.C.
12. Cortes, C., and V. Vapnik. 1995. Support-vector networks. *Machine Learning* 20 (3): 273–297.
13. Bishop, C. M. 2006. *Pattern recognition and machine learning*. Springer, Berlin.
14. Rish, I. 2001, August. An empirical study of the naive Bayes classifier. In *IJCAI 2001 Workshop on Empirical Methods in Artificial Intelligence*. IBM New York.
15. Breiman, L., et al. 1984. *Classification and regression trees*. CRC press.
16. McLachlan, G., K. A. Do, and C. Ambroise. 2005. *Analyzing microarray gene expression data*. John Wiley & Sons.
17. Schapire, R.E. 1990. The strength of weak learnability. *Machine Learning* 5 (2): 197–227.

Mobile Internet+ Campus Bicycle Sharing System Planning

Lisha Shi and Lan Wu

Abstract Campus bicycle sharing system is a new form of transportation resource sharing, which builds an App-based bike leasing platform between bike owners and users, while reserving the teachers' and students' ownership of their bicycles. The system, based on analysis of leasing demand, optimizes planning and setting of bicycle sharing stations to be put into service. Mobile Internet technology and electronic dynamic password lock are applied for self-service leasing business at client.

Keywords Bicycle sharing · Bus rapid transit · Sharing stations planning

Introduction

Started in the 1960s and evolved for nearly half a century, bicycle sharing system currently has become robust and will be growing and improving in China [1]. With constantly expanding construction scale of university campuses in recent years, functional zones of campus are located farther from each other and far beyond the range of travel on foot. As an environment-friendly, easy-to-use, and low-speed transportation, bicycle is suitable for use on path of campus, and more and more teachers and students tend to travel within campus by bike instead of on foot. However, due to long traveling distance, centralized traveling hours and huge number of the users, popularity of bicycles cause some problems which are getting serious. Firstly, insufficient parking areas and disorderly parking condition result in low use of bikes and more than 80% of the bicycles are left idle. Secondly, theft of bike often happens in campus, and thereby illegal trading on black market is getting intensive. Management measures similar to those applicable to public bikes in cities are taken on trial in many universities, but often encounter bottleneck during implementation. Lots of facilities are required for post-locked public bikes, which occupy massive area

L. Shi · L. Wu (✉)
School of Automobile and Traffic Engineering, Nanjing Forestry University,
Nanjing 210037, China
e-mail: wulan_nj@163.com

and consume huge workload, need long-term high investment in hardware equipment and system software, and thereby complex and miscellaneous system and equipment maintenance activities. In addition, unbalance in bike allocation and inconvenience in scheduling causes difficulties in borrowing and returning bike. Therefore, a new and applicable form of bicycle sharing solution is in demand.

This campus bicycle sharing system encourages teachers and students to put their own bikes into the sharing cycle, optimizes planning and setting of sharing stations within the area, uses App system to provide a leasing platform for the bike owners and temporary users, and thereby achieves an intelligent self-service borrow-and-return process.

The General Design of Campus Bicycle Sharing System

Design Concept

This campus bicycle sharing system is designed based on research and analysis of Nanjing Forestry University, as well as mobile Internet technology, integrating smartphone OS, and electronic dynamic password lock. On the basis of preliminary planning within campus, existing bicycles and intensive resources are utilized to provide safe, quick and convenient, and environment-friendly travel transportation, so as to promote the teachers and students to join this sharing program.

In the context of campus, attracting idle bikes to the sharing cycle will promote utilization of transportation resources, reduce travel cost, provide a new travel option and boost development of harmonious campus.

System Structure

This system serves bicycle owners, temporary users, and system administrators. The system consists of parking subsystem and allocation subsystem for bike owners, borrowing subsystem and changing subsystem for temporary users, and database management system and payment subsystem at the management center. Once a bike owner decides the parking and sharing location and standby duration, the available bike is brought into the sharing cycle, and will be used by a temporary user via App in the system, which is studied by López et al. [2]. A sort of dynamic password lock which is allocated consistently will be used to lock the bicycles standby, which will be unlocked when dynamic password is received via App, so as to achieve self-service sharing cycle without personal takeover. Additionally, the management center will develop effective management measures concerning admission mechanism, application zone, dynamic password lock, and system vehicle to ensure stable operation of the system. The sharing solution is illustrated as in Fig. 1.
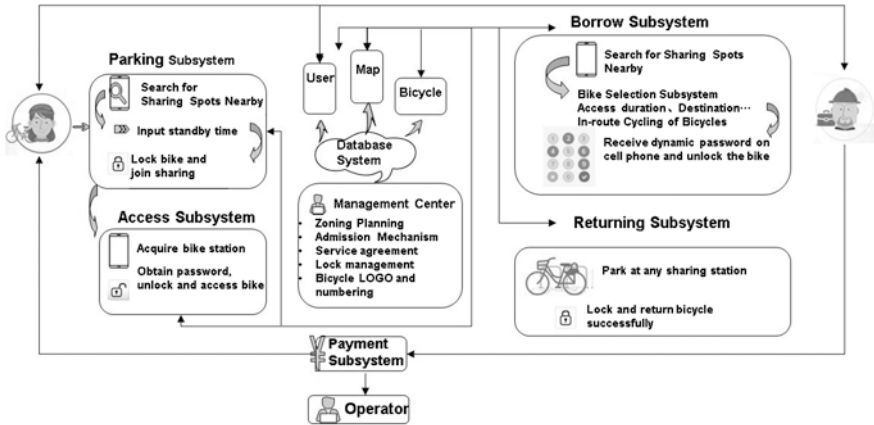
Fig. 1 Framework of the school bicycle sharing system

System Component Description

Management Protocol

To optimize management of system users and to provide efficient, safe, and easy and quick bicycle sharing service, the following management measures will be taken in design and operation of the system.

(1) Users will be registered with Campus E-Card or a user card issued by the management center to bind with the bicycle sharing system, which ensures system reliability.
(2) Bike owners will use an intelligent password lock allocated by the system upon taking part in it, to build an association between App system and operation of dynamic password locking and thereby achieve intelligent control of the lock.
(3) Before being shared, any bicycle will be attached with a logo and numbered properly for easy identification.

Planning of Sharing Stations

(1) Sharing stations planning principle

- Based on local condition in the zone: Set up bicycle sharing stations based on diverse strategies, taking into consideration layout features of teaching zone, lodging and office area, and public traffic condition nearby.
- Convenient access and easy use: The bicycle sharing stations should be located as close to exit/entry of building and access to zones as possible within the campus, so as to give full play to door-to-door transportation service in bicycle sharing system.

- Compatibility with buildings: The sharing stations should be located in compatibility with buildings and accesses, such as at path in good condition or on square accessible, so as to minimize interference with public traffic of pedestrians.
- Landscaping and environmental coordination: The sharing stations should be set up in coordination with the surroundings and in alignment with campus culture, in such a uniform way to beautify the campus and represent aesthetic features of campus [3].

(2) Zoning

The application zone will be divided into several bike transportation zones, and under an overall control based on general planning, to maximize advantages of bicycle in short-distance travel and to make it a leading way of travel within and around the campus. In planning and design of bicycle sharing stations, the campus will be zoned by block and various traffic sources will be integrated, so as to have a full knowledge of these sources and the traffic flow between them. In zoning, number of the blocks should be minimized, while the traffic sources and flows are accommodated exactly and adequately. These blocks should be homogeneous, and preferably separated with natural border such as stream [4].

The campus is now divided into ten zones based on physical space of campus, bicycle traveling routes and parking conditions, and in consideration of location and traffic conditions of functional zones within the campus, as shown in Fig. 2.

(3) Sharing stations setting

As indicated in Fig. 3, there is a highest traffic flow from student lodging area to teaching area, and from Zones 1, 7 and 9 to Zones 2 and 4, and travel along these routes will be met adequately in planning of sharing stations. In design of sharing stations, the stations are arranged and set zone by zone based on zoning of campus,

Fig. 2 Campus traffic zones division diagram

Fig. 3 Desire line
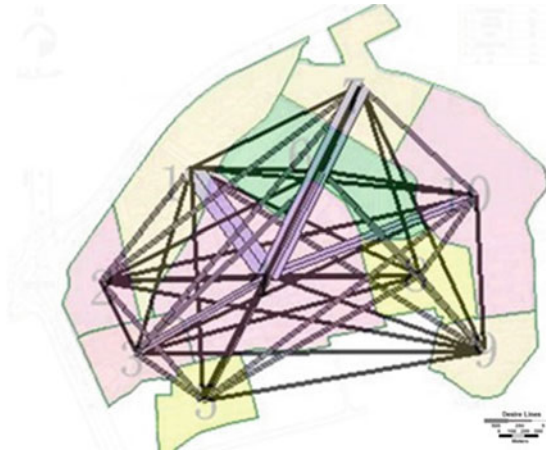
Fig. 4 Bicycle sharing points setting and distribution

current status and layout of buildings, roads and public facilities, and the traffic flow as shown in the expected route map, so as to ensure coordinated development of the sharing stations across the campus to meet the campus traffic demands. In consideration of the aforesaid factors, 25 sharing stations are set up in early stage, as shown in Fig. 4.

(4) Analysis on sharing stations setting

To provide the users with easy bicycle borrowing/returning service is one of the key functions of the sharing system, so the distance from user's location to the sharing station and from the sharing station to user's destination should be as short as practical. Bus stop/station service coverage is introduced into design of the campus bicycle sharing system. Coverage of the sharing station means the percentage of area of the circle drawn around the sharing station to the total area of campus.
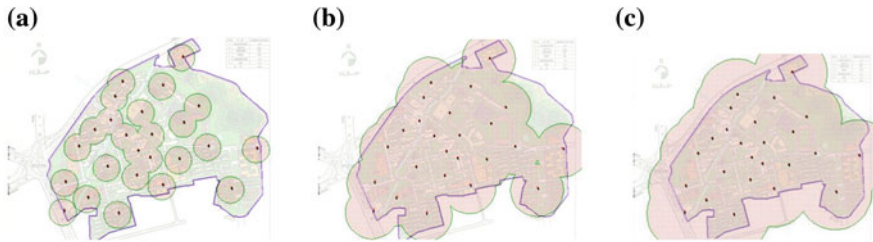
Fig. 5 **a** Walking circle of 1 min. **b** Walking circle of 2 min. **c** Walking circle of 3 min

Fig. 6 Model of dynamic
password lock
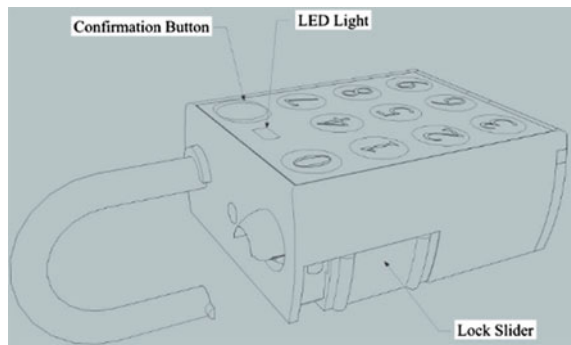
According to a survey, travelers hoping to reach a sharing station within 1, 2, 3, 4, or 5 min, respectively, account for 15, 40, 24, 12, and 9% of the total respondents. Assuming that an adult walks at a speed of about 1.2 m/s, the distances of 1/2/3 min' walk are, respectively, 72, 144, and 216 m, as shown in Fig. 5 and with coverage of about 42, 95, and 99%, indicating that this setting plan provides good accessibility and convenience.

Design of Intelligent Dynamic Password Lock

The bikes will be locked with a sort of electronic dynamic password lock which is associated with App and key-free, and will be unlocked when a valid password (which is changed dynamically) is entered via keypad on the lock.

This application is to research design of a dynamic password lock based on SCM and GSM module. A model of electronic dynamic password lock is illustrated as in Fig. 6. A bike owner parks his/her bicycle at a specified sharing station, logs to the system to identify the station, and locks the bike to bring it into the sharing cycle. When a temporary user identifies lease of a bike in the system, a dynamic password is generated from a random number by SCM and sent to the specified cell phone via GSM module [5, 6]. The user acquires and enters the password via keypad on the

lock to unlock the bicycle and so gets access to it. As required specially by the bicycle sharing system, when the temporary user parks the bike during use, he/she can lock it without log-on to the system, and then unlock it by using the original password acquired when getting access to it in an easy but safe way.

Client Application Design

The application client (App) is a software operating on cell phone terminal generally based on Windows mobile, iPhone, BlackBerry, or Android OS. Here Android OS is taken, for example, in design of application client for the campus bicycle sharing system.

The system software App mainly contains the following business logics, as shown in Fig. 7.

(1) User's registration and log-on.
(2) Bike owner selects a sharing station, available period, and then parks his/her bicycle and joins it to the sharing cycle, which will be borrowed by a temporary users via the system.
(3) A temporary user searches for the sharing stations around him/her by inputting use time and destination or related information, and the system will find the available
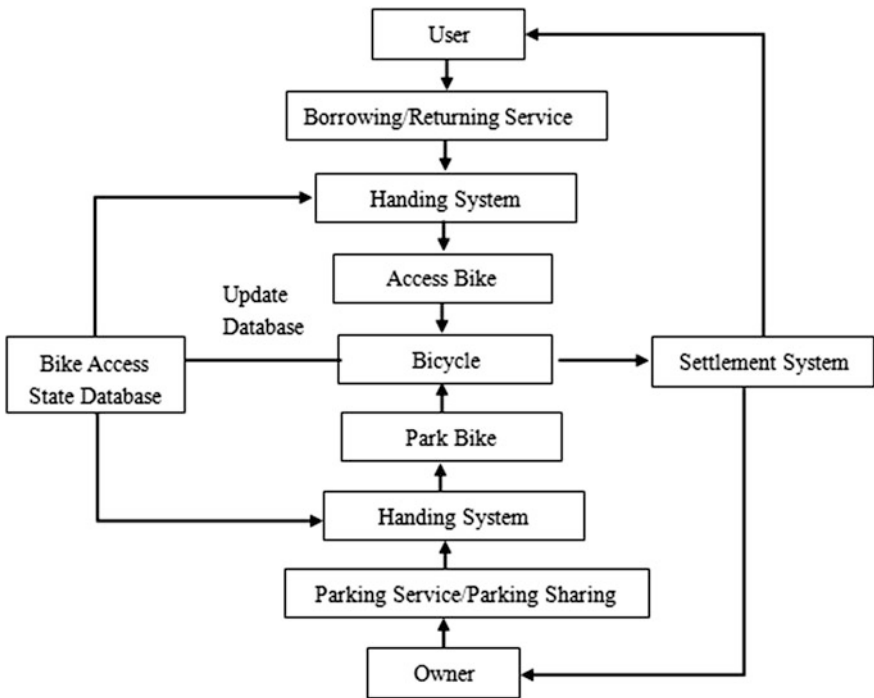
Fig. 7 System Logic

bikes and recommends those which are parked at the destination to achieve in-route cycling of bicycles through travels. The user selects a bike to borrow and acquires the dynamic password to unlock and accesses it. Any user who has borrowed a bike but not returned it will not be allowed to borrow another bike.

(4) A user who has finished using a bike borrowed can park it at any sharing station which he/she selects in the system and locks it to finish returning. And the dynamic password will be reset. Payment will be made via Alipay or e-bank or a third-party payment platform. If a user wants to park the bike halfway, he/she can lock it directly without log-on to the system, and then unlock it by using the original password.

(5) A bike owner can locate his/her bicycle upon withdrawing it and acquires the unlocking password and then withdraws it successfully. If the specified available period has expired but the bicycle has not been returned, the system will urge the borrower by sending a short message.

Application Prospect

This bicycle sharing system can be introduced not only to campus of university but also to community, industrial park zone, and commercial district, to provide easy and quick door-to-door travel service for users when scheduled properly and even expanded toward the whole of city. It is advisable to integrating bicycle sharing with public transit in sustainable urban transportation framework.

In terms of application technology, instant user messaging module can be integrated during bicycle sharing to enable real-time communication between the users throughout sharing. XMPP which is currently being used widely is a cross-platform instant messaging protocol and can be applied in development of the client to enable adding friend or transmission of text, image, or voice, making the system more user-friendly.

In addition, the system can be connected to technical service system of cloud computing or Big Data, providing added value in bicycle circulating data. Large-scale foundational platform which is closely integrated with bicycle sharing system can be built by using cloud computing to perform data storage and analysis. Bicycle circulating data in the system can provide a full and objective basis for growth and improvement of the sharing system, and provide a data platform for development of bicycle traffic or even the whole transportation system.

Conclusion

The new-type bicycle sharing system integrates user's own bikes into the system, instead of centrally manufacturing and putting public bicycle into use, thus enabling recycling of existing bikes within campus, solving the problem of parking, and

effectively integrating value of bikes which are idle in campus. Setting fixed parking lots based on preliminary planning and traffic conditions of application zones for the sharing system facilitates centralized parking and management and thereby enhances security. In solving bicycle traffic problem within campus, use of App platform and dynamic password lock reduces civil engineering. Combination of mobile Internet + and smartphone abandons conventional service terminal with LCD screen and interacting keys and lockable post.

Acknowledgment

References

1. Zhang, R., X. Wu, and L. Lei. 2011. Research on the Development and Strategy of Bike Sharing System in Urban Transportation. *Journal of Transportation Systems Engineering and Information Technology* 9 (4): 20–26.
2. López, L., E. María, and D.A. Monzón. 2014. Shared Bike Systems: Intelligent Vehicles for Sustainable Cities. *Carreteras* 4 (194): 80–88.
3. Geng, X., K. Tian, Y. Zhang, and Q. Li. 2009. Bike Rental Station Planning and Design in Paris. *Urban Transportation* 7 (4): 21–28.
4. Yao, Y., and Y. Zhou. 2009. Bike Sharing Planning System in Hangzhou. *Urban Transportation* 7 (4): 21–28.
5. Dong, Y., and D. Li. 2014. The Design of Dynamic Password Lock Based on Singlechip and GSM. *Computer CD Software and Application* 16: 65–67.
6. Qiu, J., et al. 2015. Comprehensive Assessment of Urban Transportation Using Big Data. *Urban Transportation* 13 (3): 63–70.

Analysis of the Homogeneity of Driver Behaviors at Intersections

Le-yi Wang, Xiao-bei Jiang, Qian Cheng, Cheng-gang Li
and Wu-hong Wang

Abstract Intersections are road traffic accident-prone locations. About 10–40% of the traffic accidents take place at intersections worldwide every year. An intersection in Haidian District, Beijing was surveyed by video survey method, and the survey videos were processed and analyzed. The velocities and trajectories of the vehicles at the intersection were collected. Analysis of the homogeneity of driver behaviors shows that vehicles have the best homogeneity of velocities when passing the stop line and the crosswalk. Right-turning vehicles have the best homogeneity of velocities, left-turning vehicles take the second place, and homogeneity of velocities of straight-going vehicles is the worst. Straight-going vehicles have the best homogeneity of trajectories, right-turning vehicles take the second place, and left-turning vehicles have the worst homogeneity of trajectories. The results provide references for driver behavior modeling at intersections and the development of intersection driving assistant systems.

Keywords Driver behavior · Homogeneity · Intersection · Traffic safety

Introduction

Traffic safety is one of the major problems faced by the road traffic nowadays. Safety of human life and property is seriously threatened by traffic accidents all around the world. Intersections are accident-prone locations. About 10–40% of the traffic accidents take place at intersections all over the world every year [1]. It is estimated that about 65% of the traffic accidents are only caused by driver factor, and about 95% are related to drivers [2], which indicates that driver is the most important influence factor of road traffic safety.

L. Wang · X. Jiang (✉) · Q. Cheng · C. Li · W. Wang (✉)
School of Mechanical Engineering, Beijing Institute of Technology, Beijing 100081, China
e-mail: jiangxiaobei@bit.edu.cn

W. Wang
e-mail: wangwuhong@bit.edu.cn

Many researches focused on driver behaviors at intersections for the past few years. Zhang et al. [3] studied drivers' decision-making behaviors at non-signalized intersections based on dynamic reduplicate game theory, and the results show that impulsive drivers tend to choose acceleration as the first strategy; mild drivers prefer acceleration or uniform strategy; cautious drivers prefer uniform or deceleration strategy. A study by Jin et al. [4] shows that sight obstruction will decrease the gap acceptance of right-turning drivers at intersections, which causes low capacity, delay, and more potential conflicts between right-turning vehicles and pedestrians. A study by Werneke and Vollrath [5] shows that drivers' attention distribution and behaviors when entering an intersection are related to the environmental characteristics (including the density of the oncoming traffic from the left and the number of the chunks of information related to the driving task) of the intersection. A study on driver behaviors at signalized intersection by Rittger et al. [6] shows that drivers choose a larger acceleration when the red light turns to green compared to when the traffic light is solid green and drivers decelerate earlier when the traffic light is solid red compared to when the green light turns to red. Danaf et al. [7] compared the differences in driver behaviors and red light running between different countries and proved that traffic violations are not necessarily associated with dangerous driver behaviors, but with drivers' subjective intentions to violate traffic rules.

Theeuwes and Godthelp [8] put forward the concept of self-explaining road (SER), which is defined as a traffic environment, which elicits safe behaviors simply by its design. A study by Van Vliet and Schermers [9] proposed that self-explaining road design method focuses on three key principles of functionality, homogeneity, and predictability. Homogeneity requires that physical appearances of roads within one road category should be homogeneous, so that road users' behaviors on that category of roads will be homogeneous. Charlton et al. [10] finished a self-explaining road reconstruction project. After reconstruction, speeds on local roads decreased significantly and homogeneity of speeds on both local roads and collector roads increased significantly. The project also created a 30% reduction in the number of accidents and an 86% reduction in the damage of accidents per year. A study by Mackie et al. [11] shows that self-explaining road reconstruction created obvious differences in road users' behaviors between roads from different categories and increased the homogeneity of road user behaviors on roads within the same road category.

The paper surveyed the Wanliu Middle Road–Wanquanzhuang Road intersection in Haidian District, Beijing, made image coordinate-world coordinate transformation of the survey videos and collected the data of velocities and trajectories of the vehicles. The homogeneity of driver behaviors was analyzed. The research provides references for driver behavior modeling at intersections, intersection safety design, and the development of intersection driving assistant systems.

Methods

Basic Information of the Intersection Surveyed

The paper surveyed the Wanliu Middle Road–Wanquanzhuang Road intersection in Haidian District, Beijing. Location of the intersection is shown in Fig. 1 and a photo of the intersection is shown in Fig. 2. The intersection is signalized and the two roads connected by it are both two-way, four-lane roads with central median. The main land use type around the intersection is residential land. Complete the investigation in off-peak hours in order to avoid the influence of the high traffic volume on driver behavior analysis. Choose video survey method and use a video

Fig. 1 Location of the intersection surveyed

Fig. 2 A photo of the intersection surveyed
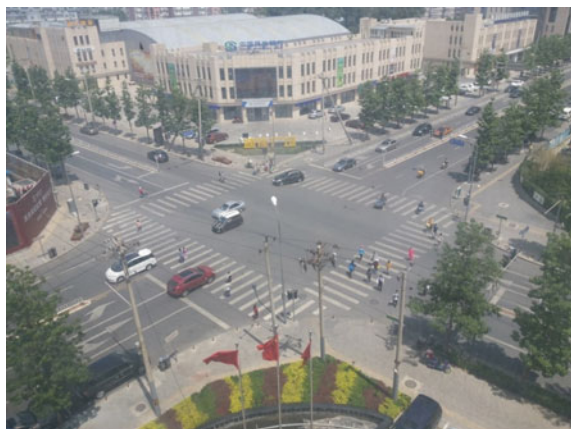
Table 1 Traffic volume at the intersection

Direction		Type of road users			
		Lorry/bus	Car	Motorcycle	Non-motor vehicle
North	U-turning	0	4	0	0
	Left-turning	2	82	8	25
	Straight-going	37	222	10	33
	Right-turning	0	39	0	8
South	U-turning	0	33	0	0
	Left-turning	0	23	2	6
	Straight-going	39	197	25	25
	Right-turning	0	74	19	31
East	U-turning	0	16	0	0
	Left-turning	2	64	25	27
	Straight-going	0	123	21	39
	Right-turning	2	148	6	21
West	U-turning	0	6	0	0
	Left-turning	0	58	2	6
	Straight-going	0	99	21	31
	Right-turning	0	70	8	4

Table 2 Pedestrian volume at the intersection

Direction	The left crosswalk	The right crosswalk	Sum
North → South	93	140	232
South → North	280	111	391
East → West	228	99	327
West → East	105	148	253

camera and a tripod as the equipment. Traffic volume during the survey period at the intersection is shown in Tables 1 and 2.

Data Processing and Analysis

Make image coordinate-world coordinate transformation of the survey videos. The simplified transformation model is shown as formula (1) [12]:

$$X = \frac{c_1 x + c_2 y + c_3}{c_4 x + c_5 y + 1}$$
$$Y = \frac{c_6 x + c_7 y + c_8}{c_4 x + c_5 y + 1} \tag{1}$$

(a)   **(b)**

The entrance lane group   the intersection area group
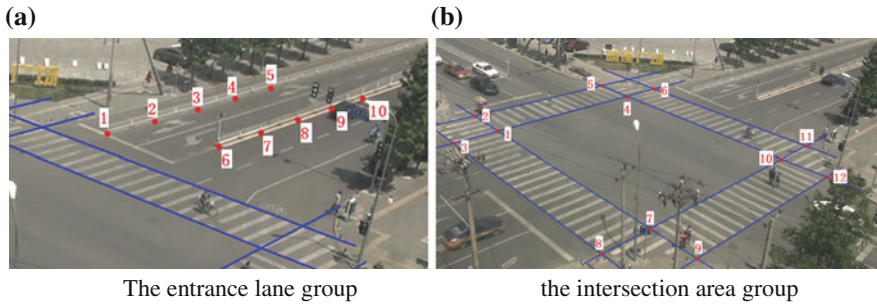
Fig. 3 Ground reference points

In which:

X, Y world coordinate (cm)
x, y image coordinate (pixel)
$c_1, c_2, \ldots, c_8$ transformation coefficients

The coordinate transformation model has eight unknown parameters, which requires at least four ground reference points to solve it. For each point on the image, using the reference points near it to fit the transformation model can reduce the error of the coordinate transformation. So two group of ground reference points are used: the entrance lane group (including 10 reference points), as shown in Fig. 3a and the intersection area group (including 12 reference points), as shown in Fig. 3b. The two groups are used for coordinate transformation of points on the entrance lane and the intersection area, respectively.

Using the data of image coordinates and world coordinates of the ground reference points to make quadratic nonlinear fitting of the coordinate transformation formula by MATLAB. To extract the velocities and trajectories of the vehicles from the survey videos, set a crucial frame at every 15 frames (0.6 s) in the survey video, record the image coordinate of the projective point on the ground of the midpoint of the front bumper of the vehicle as the image coordinate (x, y) of the subject vehicle in each crucial frame. Use the coordinate transformation formula after fitting to compute the world coordinates (X, Y) of the subject vehicle in each crucial frame, and further obtain the velocity and trajectory of the subject vehicle.

Analysis of Driver Behaviors

Analysis of Vehicle Velocities

Velocities of left-turning, straight-going, and right-turning vehicles are shown in Fig. 4.

Fig. 4 Vehicle velocities

(a)

Velocities of left-turning vehicles

(b)

Velocities of straight-going vehicles

(c)

Velocities of right-turning vehicles
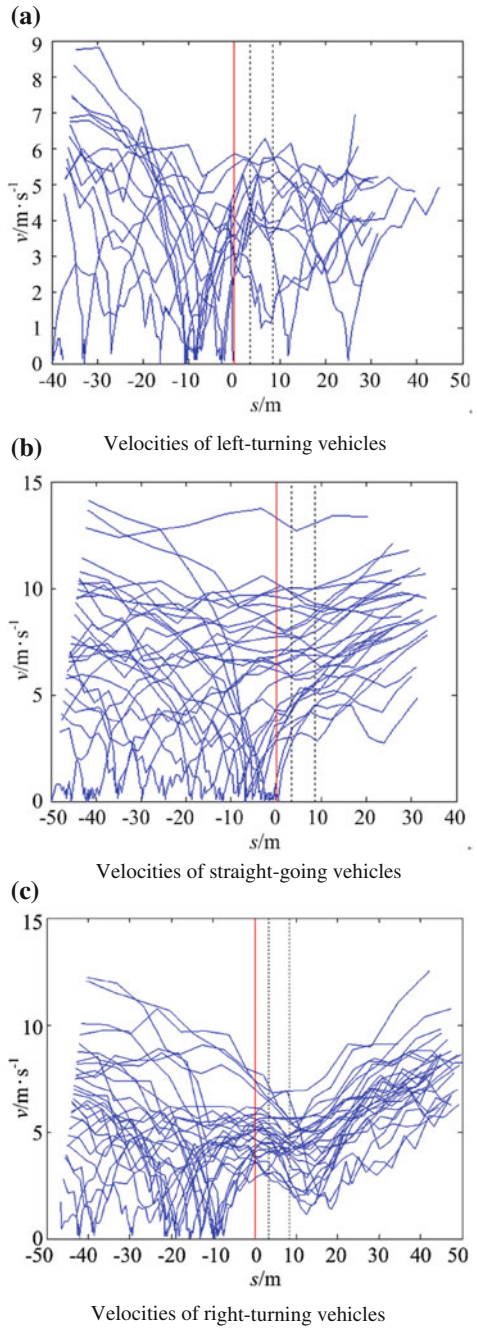
(In Fig. 4, *x*-axis is the distance from the position of the vehicle to the stop line; *y*-axis is the velocity of the vehicle at that position; the straight solid line represents the stop line, and the straight dash lines represent the boundaries of the crosswalk.)

It can be concluded that:

(1) Left-turning vehicles with a speed higher than 6 m/s have an obvious deceleration about 30 m before the stop line. The left-turning vehicles near the stop line or having passed the stop line usually keep a speed lower than 6 m/s. If there are no other road users at the intersection, free speed of the left-turning vehicles to get through the intersection will be about 6 m/s. Only a few left-turning drivers tend to slow down when approaching the crosswalk.

(2) Before passing the stop line, the main reason for the straight-going vehicles to decelerate is the influence of the signal lights and the leading vehicles. If there are no restrictions, straight-going vehicles tend to get through the intersection with a speed of more than 10 m/s. Only a few straight-going drivers slow down when approaching the crosswalk, but the homogeneity of the velocities get better near the crosswalk. Straight-going vehicles will accelerate immediately after they pass the crosswalk. Some vehicles have a high free speed of about 13 m/s.

(3) Right-turning vehicles with a speed higher than 7 m/s will decelerate before the stop line and pass the stop line with a speed of about 7 m/s. The homogeneity of velocities of the right-turning vehicles increased significantly when passing the stop line and the crosswalk.

(4) Left-turning, straight-going and right-turning vehicles all have bad homogeneity of velocities before the stop line, because of the influence of signal lights. They have the best homogeneity of velocities when passing the stop line and the crosswalk. Right-turning vehicles have the best homogeneity of velocities, left-turning vehicles take the second place, and straight-going vehicles have the worst homogeneity of velocities.

Analysis of the Homogeneity of Vehicle Trajectories

Trajectories of the left-turning, straight-going and right-turning vehicles are shown in Fig. 5.

(In Fig. 5, the curves represent the trajectories of the vehicles; *x*-axis and *y*-axis are the world coordinates of *X*/m and *Y*/m; the straight solid line represents the stop line, and the straight dash lines represent the boundaries of the crosswalk.)

It can be concluded that:

(1) Almost all of the left-turning vehicles complete lane changing 15 m before the stop line. They keep a good homogeneity of trajectories from 30 to 10 m before the stop line and have bad homogeneity of trajectories from passing the stop line to completing turning left.

Fig. 5 Vehicle trajectories

(a)

Trajectories of left-turning vehicles

(b)

Trajectories of straight-going vehicles

(c)

Trajectories of right-turning vehicles
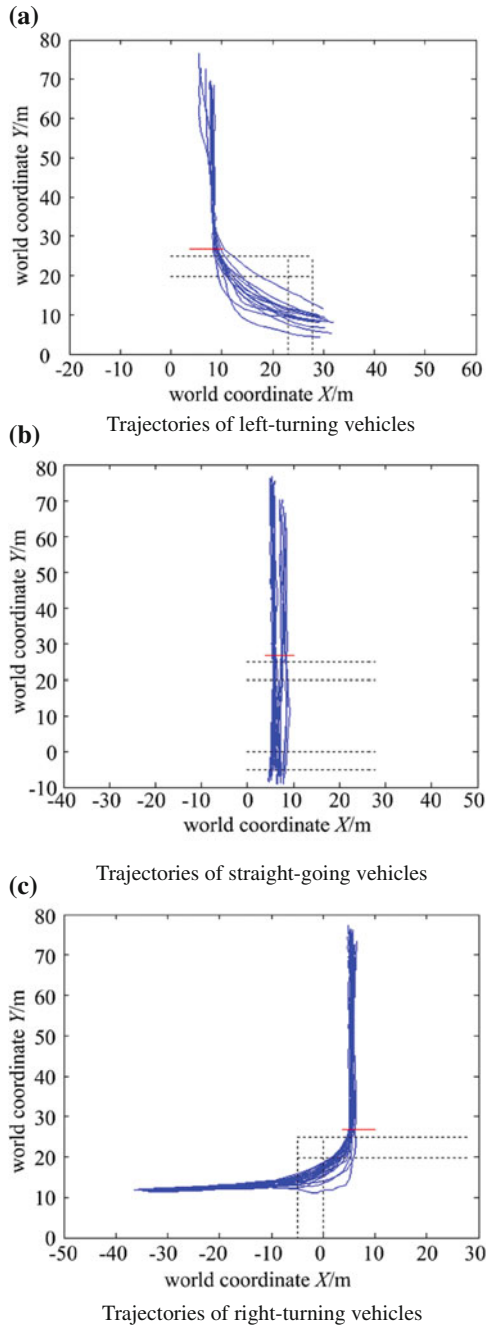
(2) Few straight-going vehicles change their lane before the stop line. Straight-going vehicles keep a good homogeneity of trajectories during the whole process of getting through the intersection.

(3) Few right-turning vehicles change their lane. Right-turning vehicles keep a good homogeneity of trajectories before they pass the stop line and have bad homogeneity of trajectories from passing the stop line to completing turning right.

(4) In general, straight-going vehicles have the best homogeneity of trajectories, right-turning vehicles take the second place, and left-turning vehicles have the worst homogeneity of trajectories.

(5) The region, where the trajectories of the left-turning, straight-going and right-turning vehicles and the crosswalk overlap with each other, forms the left-turning, straight-going and right-turning vehicle–pedestrian conflict danger zone. The area of the left-turning vehicle–pedestrian conflict danger zone is the largest, because of the bad homogeneity of trajectories of the left-turning vehicles. The trajectories of left-turning vehicles have large variation, which causes potential accident risk. In intersection safety design, using methods such as channelization island and road markings to increase the homogeneity of vehicle trajectories can reduce the area of the vehicle–pedestrian conflict danger zone and improve intersection traffic safety.

Conclusions

This paper surveyed the Wanliu Middle Road–Wanquanzhuang Road intersection in Haidian District, Beijing using video survey method, made image coordinate-world coordinate transformation of the survey videos and extracted the velocities and trajectories of the vehicles from the survey videos. Analysis of the homogeneity of driver behaviors comes to the following conclusions:

(1) All of the vehicles show the best homogeneity of velocities when passing the stop line and the crosswalk. Right-turning vehicles have the best homogeneity of velocities, left-turning vehicles take the second place, and straight-going vehicles have the worst homogeneity of velocities.

(2) Straight-going vehicles have the best homogeneity of trajectories, right-turning vehicles take the second place, and the homogeneity of trajectories of the left-turning vehicles is the worst. Left-turning and right-turning vehicles have the worst homogeneity of trajectories when they are turning.

(3) Bad homogeneity of trajectories of the vehicles will enlarge the area of vehicle–pedestrian conflict danger zone, thus affecting intersection traffic safety.

Results of this paper provide references for driver behavior modeling at intersections, intersection safety design and the development of intersection driving assistant systems.

References

1. Zou, H.B., and J.M. Xu. 2008. Application of traffic conflict technology in assessing the intersection safety in Guilin. *Western China Communications Science & Technology* 1: 20–22.
2. Wang, W.H., H.W. Guo, and W.W. Guo. 2013. *Transportation behavior analysis and traffic safety evaluation*, 54–55. Beijing: Beijing Institute of Technology Press.
3. Zhang, L., X.W. Huang, and W.M. Wu. 2014. The analysis of driver's behavior in non-signalized intersection based on the game. *Applied Mechanics and Materials* 505–506: 1157–1162.
4. Jin, J., W. Wang, G. Wets, et al. 2014. Effect of restricted sight on right-turn driver behavior with pedestrians at signalized intersection. *Advances in Mechanical Engineering* 2014 (1): 1–6.
5. Werneke, J., and M. Vollrath. 2014. How do environmental characteristics at intersections change in their relevance for drivers before entering an intersection: analysis of drivers' gaze and driving behavior in a driving simulator study. *Cognition, Technology & Work* 16 (2): 157–169.
6. Rittger, L., G. Schmidt, C. Maag, et al. 2015. Driving behaviour at traffic light intersections. *Cognition, Technology & Work* 17 (4): 1–13.
7. Danaf, M., S. H. Hamdar, M. Abouzeid, et al. 2016. Comparative assessment of driving behavior at signalized intersections using driving simulators. *Journal of Transportation Safety & Security* 1–35.
8. Theeuwes, J., and H. Godthelp. 1995. Self-explaining roads. *Safety Science* 19 (2–3): 217–225.
9. Van Vliet, P., and G. Schermers. 2000. Sustainable safety in the Netherlands—A new approach. In *Traffic Calming from Analysis to Solutions—Proceeding of the Extraordinary Workshop of ICTCT*, New Delhi, India, March 2000. Austria: ICTCT Secretariat, 10–27.
10. Charlton, S.G., H.W. Mackie, P.H. Baas, et al. 2010. Using endemic road features to create self-explaining roads and reduce vehicle speeds. *Accident Analysis and Prevention* 42 (6): 1989–1998.
11. Mackie, H.W., S.G. Charlton, P.H. Baas, et al. 2013. Road user behavior changes following a self-explaining roads intervention. *Accident Analysis and Prevention* 50 (1): 742–750.
12. Li, J., C. Shao, A. Haghani, et al. 2012. Video-based pedestrian traffic parameters extraction. Transportation Research Board 91st Annual Meeting.

Comparative Analysis of the Public Transit Modes Based on Urban Area Location Theory

Lan Wu

Abstract Different characteristics of transportation supply and demand in the big cities and small or medium-sized cities determine different public transportation modes supported the TOD. Many new districts are becoming the main location of urban spatial expansion, and also the most important areas in the process of urbanization. Planning and construction of new district have become the important means and opportunity to restructure the city space. TOD concept is easier to achieve in the new district. Based on the location character of new district, from the perspective of work and living balancing, the suitability was analyzed between rail transit and BRT. Utility analysis was adopted to study the main factors of affecting BRT and conventional public transit utility value. By comparison, BRT is more suitable as the backbone to orient new district development.

Keywords New district · Transit-oriented development · Bus rapid transit · Utility

Introduction

Public transit system is mainly made up of rail transit (metro and light rail), bus rapid transit, and bus lines, which support TOD (traffic oriented development). The TOD is one of the most representative models of new urbanism, and it is applied in the urban development broadly, especially in new districts. In the city, the TOD consists of "new built-up area TOD," "redevelopment area TOD," and "need-to-be developed area TOD." In the construction of TOD project, the new built-up area TOD is apt to realize, while the planning and design of TOD in the redevelopment area face many problems such as complicated transport, land use being hard to change, and financial matter in the built-up area. In addition, urban

L. Wu (✉)
School of Automobile and Traffic Engineering, Nanjing Forestry University, Nanjing 210037, China
e-mail: wulan_nj@163.com

morphology is relatively stable in the built-up area, so implement of TOD will not affect city's evolution dramatically.

New district is a new built area in the process of cities evolution and development, and it can reduce employment and living pressure of main area of cities. The new district is an important part of cities but independent of main area. Because TOD model is easily to be realized in the new district, research will be based on the new district type construction. Comparative analysis of oriented transportation modes in the new district development will be conducted.

Definition of New District

There are six types of new districts roughly based on the different industry dominant models or structures: administration new district, university town, tourism new district, science new district, sports new district, and industry new district. From the perspective of how new districts come into being and development, new districts of cities are the result of governments' plan, guidance, and regulation, of which the purposes are to disperse city population, improve city environment, transfer or add city function, etc. Objectively, there is not absolute boundary between the new district and central urban area [1].

According to the practice of new cities and urbanization development in China, new district is defined as: With the need of city spatial development, the new district is a centralized region with certain scale and density in the process of urbanization, which is newly built after overall plan in fringe areas with certain distance from central urban area and has certain functions. Essentially, the new district is a kind of comprehensive development of city functions. In form, it is a relatively independent city region geographically and is attached to central urban area strongly; the new district is very different form from new city. The new city is larger compared with the new district in scale and has city size; compared with old city, distance from the new city to the old city is longer than the distance from the new district and the central urban area, and there is a clear boundary between the new city and the old city.

Comparison Among Public Transit Modes Oriented in the New District

Comparison Between BRT and Rail Transit

Public transit modes in the city development are mainly made up of rail transit (metro and light rail), bus rapid transit, and bus lines, the scope of application for three kind of public transit and pedestrian traffic is shown in Fig. 1. The rail transit

Fig. 1 Scope of application for urban passenger transport
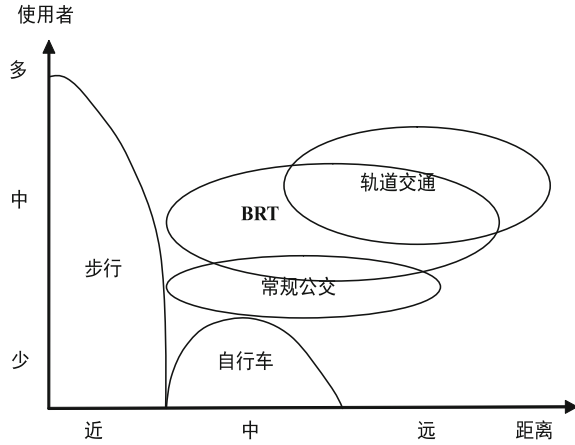
is the backbone transit in big city because of its high operation efficiency and fixed operation time, it fits long-distance transport connection in the big cities, even intercity transport connection. BRT can be the extension lines of rail transit for big cities such as Shanghai, Beijing, and cities with rail transit to orient new districts' development. For medium and small cities, BRT can be the backbone transit of public transit to orient urban sustainable development [2].

From Fig. 1, it can be seen that BRT and rail transit both fit medium-long distance trip. For rail transit-oriented cities, city area is always large, and new district is far away from central urban area because of cities' characteristics. It is found that construction of rail transit influences residents greatly from the perspective of commuting by investigation in Beijing. Commuting time for residents who buy houses and residents who rent houses is quite different. Rail transit will increase commuting time for residents who buy houses while decrease commuting time for residents who rent houses. As to trip distance, commuting distance of residents who live along rail transit is longer than that of residents who do not live along rail transit in most new district. In addition, commuting distance of residents who buy houses is longer than that of residents who rent houses in the district; commuting distance of residents who live in the new district is longer than that of residents who live in the central urban area. Residents along rail transit live and work in separate places, who are influenced by rail greatly, and distance between work and living is large. After the new district is well developed, employment proportion will increase gradually, and employment opportunities will be found along the rail transit lines. Industrial city in the suburb of the city, work and living is relatively independent and their spaces are evenly divided. Moreover, research in the new district shows that the distance between O-D and metro station affects trip mode a lot. In general, when distance between work place/residence and metro station is above 2 km, the proportion of choosing metro decreases sharply [3].

In 2020, all suburbs and more than half of new towns will be connected by the Shanghai Metro. TOD and TND (traditional neighborhood development) are the basic ideas of Shanghai rail transit construction, which orienting new town development. The "new town," developing with rail transit closely, can develop vertically with the guidance of rail transit; gradually, banding-group urban development structure will come into being. The Guangzhou–Foshan line of Guangzhou Metro connects Guangzhou and Foshan. Guangzhou has planned new rail transit to strengthen the connection between suburb and central urban area in 2020 and to support interconnection of the central urban area and four districts and two cities. Meanwhile, it will strengthen the connection between new and old urban areas organically and guide population of old urban area moving to new urban area.

Thus it can be seen that rail transit is used as backbone transportation model of big cities or mega cities, it can connect new districts or new towns which are relatively far from central urban area, and it can even connect two cities. Large number of residents live in new towns, where are along rail line and far from central urban area, so that commuting traffic volume become large, it is overcrowded at commuter time, while passenger flow is relatively small at other time. When passenger traffic demand from central city to new towns or another city reaches to 4 million and above, and the distance from central urban area of mega cities or big cities to new towns or new districts is relatively far away, meanwhile big cities have sufficient financial resources, rail transit can be adopted.

BRT system, combining advantages of metro and conventional bus, has low cost and high benefit, it is suitable for cities or regions to adopt, which do not have enough population but have urgent demand of passenger in the urban development. Tianjin Binhai new district has airport logistics zone, Binhai high-tech zone and development zone, which have large number of jobs. Since the bus preserved lane has been designed when planning the new district, BRT system is built to serve commuting traffic. Meanwhile, BRT demonstration line of Gangcheng Avenue in the Binhai New District will be transition and supplement of rail transit system. BRT system in Guangming New District of Shenzhen will cultivate passenger flow for long-term rail transit. BRT system in Guangzhou connects new districts such as development zone and university town and sets the record of busiest bus line of Asia in the first year of operation. Calculated by construction cost, the cost of Guangzhou BRT system is only 1/10 of 23-km metro. Qunli New District of Harbin also chooses to reserve BRT line to solve commuting problem at peak hour. To meet the traffic demand of colleges and universities, economic development zone, and industry zone's construction and development, BRT line of Zhengdong New District and new and high-tech zone has been extended. Most passengers highly praise BRT system of Binhu New District that buses are comfortable and stable. Travel time has been shortened, for example, from Weigang to Binhu headquarters, it will take 40 min by bus, but only 13 min by BRT, including time to wait for signals.
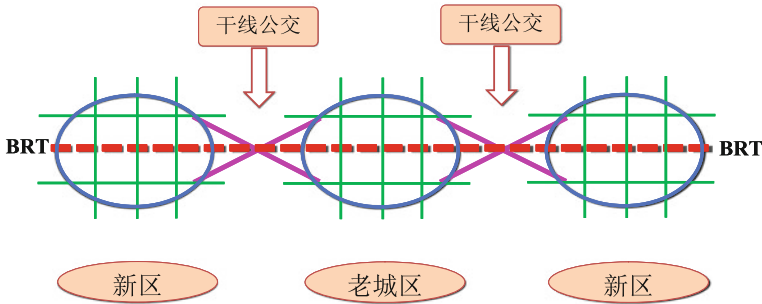
Fig. 2 Connection model between new district and old urban district

It can be concluded that BRT system can connect new districts such as new industrial zone and university town. With exclusive road right and signal priority, BRT can even has close transport capacity and speed compared with rail transit. BRT's construction and operation costs are relatively low. Existing public transport road system can be upgraded for BRT operation to a great extent. BRT system even can be a substitute model for rail transit in some cities with small-medium population. Figure 2 shows the connection model between new district and old urban district.

Comparison Between BRT and Traditional Bus

Traditional bus system is the foundation of city development, which can be applied in medium-long distance trip in the city. Traditional bus network planning should be the fundamental part in the planning of new district. Selection of large traffic volume public transport mode will be planned according to the characteristic of city development. Construction of rail transit needs to be examined and approved that the city has certain scale. In China, it is advised that when city population reaches to 2 million and above, GDP reaches to 100,000 million RMB and above, application of rail transit construction can be applied, while building BRT does not require that. Comparison between BRT and rail transit has been made, comparative analysis of BRT and other motor vehicles, such as traditional bus, taxi, and car, will be conducted next.

Trip distance is the most vital factor to influence selection of transportation modes, and trip cost (including expenditure and time), trip purpose, and traveler characteristic are also other factors. Common mode split models are include logit model [4, 5] and probit model [6]. The paper adopts multinomial logit model to analyze:

$$P_i = \frac{e^{V_i}}{\sum_{j \in A} e^{V_j}} = \frac{1}{\sum_{j \in A} e^{V_j - V_i}}, (i \in A) \tag{1}$$

In the formula, P_i is the probability of selecting ith traffic mode, V_i is the utility function of selecting ith traffic mode, and A is the set of all traffic modes. V_i function is:

$$V_i = \theta_0 + \sum_{k=1}^{n} \theta_k X_{ik} \tag{2}$$

In the formula, θ is the parameter of influence utility factor, its size shows the influence degree of influence factor; $X_i = (X_{i1}, \ldots, X_{ik})$ is the vector of influence factor. After examining factors influenced utility with χ^2 test, factors are age θ_1, income θ_2, travel time θ_3, travel cost θ_4, having bike or not θ_5, having car or not θ_6, trip purpose θ_7, etc. Define set A as the set of motorized transportation modes:

$$A = \{i = 1(\text{bus}), 2(\text{BRT}), 3(\text{taxi}), 4(\text{car})\} \tag{3}$$

Based on the different trip distances, set trip distance as short distance (8 km), medium-long distance (16 km), and long distance (24 km) then proceed regression analysis, utility function can be got from formula (2):

$$V_i = \theta_0 + \sum_{k=1}^{n} \theta_k X_{ik} = \theta_0 + \theta_1 X_{i\,\text{age}} + \theta_2 X_{i\,\text{income}} + \theta_3 X_{i\,\text{time}} + \theta_4 X_{i\,\text{cost}}$$
$$+ \theta_5 X_{i\,\text{bike}} + \theta_6 X_{i\,\text{car}} + \theta_7 X_{i\,\text{aim}} \tag{4}$$

θ_0 is the parameter of willingness to travel. Substitute model parameter calibration results into formula (4), influence to traffic mode of each factor can be obtained, see Fig. 3.

In Fig. 3, utility analysis of factors for bus, BRT, taxi, and car these four modes, have been made refined from willingness, age, income, travel time, travel cost, having bike or not, having car or not, and trip purpose. With different distances, each factor has different influence on traffic mode, and cost's influence is negative utility. In Fig. 3b, c, BRT's initial utility value is positive, while traditional bus' utility is negative, which indicates for medium-long distance trip, people prefer to BRT instead of traditional bus. Add all utility values and get the result in Fig. 4.

Figure 4 shows utility value variation trend of four transportation modes with different distances. Utility value of traditional bus and BRT decreases when trip distance increases, that is, selective probability is decreasing. It is influenced by the scope of BRT and bus lines; however, BRT has great advantages compared with traditional bus.
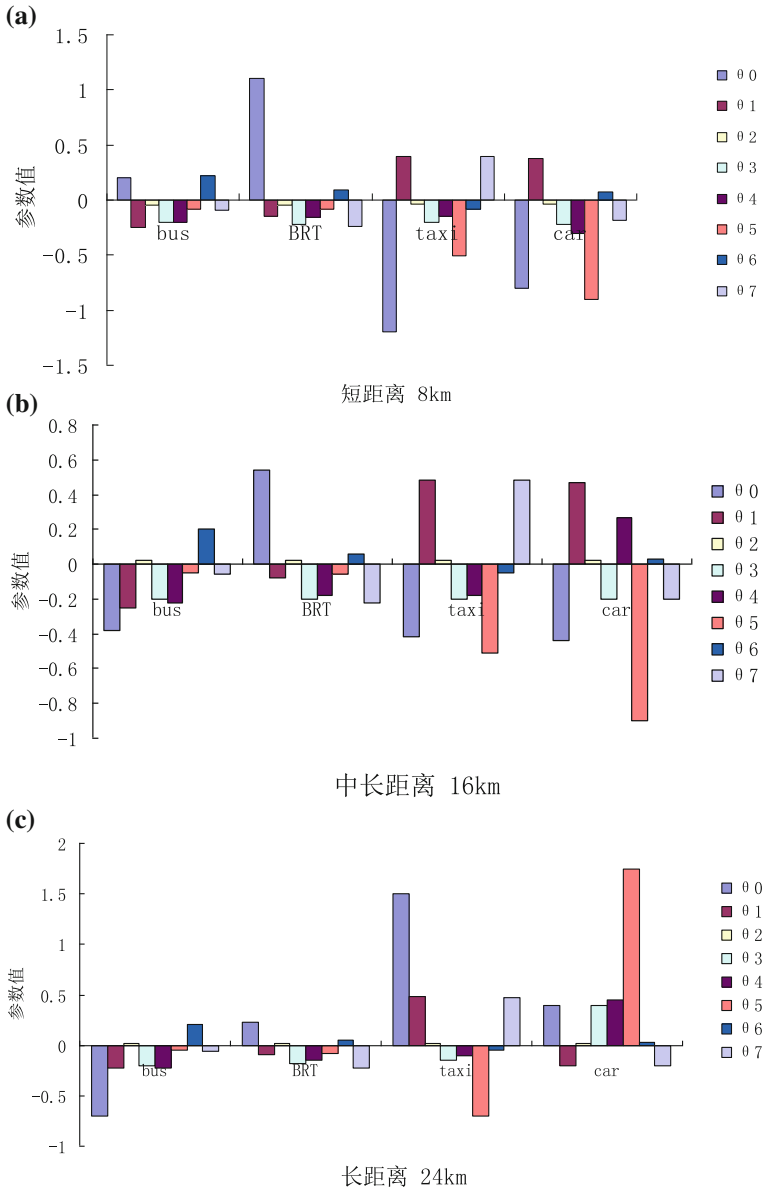
Fig. 3 θ value of each traffic mode
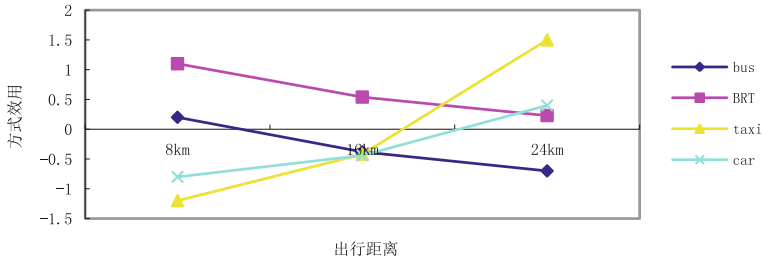
Fig. 4 Utility of four traffic modes with different distances

Conclusion

Traffic demand in new district increases with elasticity, with the development of lands and economy. New traffic demand will arise; accordingly, demand of transit facilities will increase. BRT, with characteristics of low cost, high benefit, and gradual built-up, combines metro and traditional bus' advantages and is suitable for new districts to adopt, which do not have enough population (population is increasing) but have urgent need of passenger traffic. The construction of BRT is flexible, which can be constructed by time and by space. It fits the development progress of land and is economical.

Rail transit is used as backbone transportation model of big cities or mega cities. Passenger traffic demand along rail transit should reach to 4 million person-time and above, sufficient financial resources are also required. BRT can have transport capacity for public transportation of most cities, even has close transport capacity with rail transit. Compared with traditional bus, both have low operating cost, but BRT system has higher transporting speed and better service, which reduces residents' trip cost. Considering economic rationality of operation and constraint of space, BRT system is more suitable for new districts' construction. BRT and rail transit can complement each other's advantages in a way or BRT can be rail transit's foreshadowing.

Comparison result of BRT and other motorized transportation modes shows residents are more willing to choose BRT compared with traditional bus for medium-long distance trip. By utility analysis, with the increase of trip distance, BRT has obvious advantages compared with traditional bus; therefore, BRT is more suitable for medium-long distance trip in new districts.

References

1. Chen, T., X. Liu, and X. Chen. 2009. Existing problems of urban bus-priority and its countermeasures. *Technology & Economy in Areas of Communications* 2: 72–73.

2. Wei, H., and D. Dong. 2007. The study of BRT promoting TOD in China. *Traffic and Transportation* 7: 23–25.
3. Chan, C., J. Misener, and J. Lins. 2003. Smart buses, smart intersection shine at Washington IVI meeting. *Intellimotion* 10 (3). California PATH.
4. Ben-Akiva, M., and S. Lerman. 1985. *Discrete choice analysis theory and application to predict travel demand*. Cambridge: MIT Press.
5. Domenich, T., and D. McFadden. 1996. *Urban travel demand: A behavioral analysis*. North-Holland, New York (reprinted).
6. Kumar, A. 1980. Use of incremental form of logit models in demand analysis. *Transportation Research Record* 775. TRB, National Research Council, Washington, D.C., 21–27.

Survival Analysis on Passing Time of Minor Vehicle's Road Crossing at Un-Signalized Intersection in China

Yuan Zheng, Jian Zhang, Guoqiang Zhang, Linchao Li and Biqing Ye

Abstract This paper is to develop a quantitative approach of the operation performance of left turning by means of the passing time as the indicator to analyze the interactions of traffic stream between main and minor road based on survival analysis theory at the un-signalized intersection. The data were collected by video cameras in Nanjing, China. And the whole passing time (T) was divided two sections, namely passing time 1 (T_1) in decision zone 1 and passing time 2 (T_2) in decision zone 2, respectively. A hazard-based duration model was proposed to analyze the influential variables related to the passing time, including the variables H and W with the positive effect and Q_w, S_{w1}, and L showing the negative effect on the passing time 1. The results also showed the variables revealing comprehensive characteristics are significantly negative related to the passing time 2, such as Q_e, S_{e1}, W, and C. Moreover, all variables from model estimation of T_1 and T_2 were selected into the model of whole passing time, but the variable of W has the opposite effect on passing hazard between the model of T_1 and T. In addition, the various influential variables related should be given full consideration in the

Y. Zheng · J. Zhang (✉) · L. Li
Research Center for Internet of Mobility, Southeast University,
Si Pai Lou #2, Nanjing 210096, China
e-mail: zhangjian8seu@163.com

Y. Zheng
e-mail: seuzhengy@163.com

L. Li
e-mail: 1045650765@qq.com

Y. Zheng · J. Zhang · G. Zhang · L. Li · B. Ye
Jiangsu Key Laboratory of Urban ITS, Southeast University,
Si Pai Lou #2, Nanjing 210096, China
e-mail: guoqiang.zhang@163.com

B. Ye
e-mail: 1151070746@qq.com

Y. Zheng · J. Zhang · G. Zhang · L. Li · B. Ye
Jiangsu Province Collaborative Innovation Center of Modern
Urban Traffic Technologies, Si Pai Lou #2, Nanjing 210096, China

performance analysis of left turning. It is hoped that this paper will help to enhance the safety, planning, and management of policies and facilities to improve operation efficiency at un-signalized intersection.

Introduction

Un-signalized intersections are a key element in urban streets and in rural road networks. Unlike signal-controlled intersections, un-signalized intersections provide no positive indication to the driver concerning when it is appropriate to enter the intersection. Generally, the highway un-signalized intersections are priority-controlled intersection. In principle, the mainstream has the right of priority and can drive through without stopping at the intersections, and the minor stream has to give way to the mainstream and stop appropriately. So the driver decides when to complete the required maneuver, based on elements of the decision context including distances, velocities, and vehicle performance [1].

Because the large volume of vehicles and trucks and driving priorities are not entirely respected in China, continuous traffic stream on the main and minor road often gets blocked. As a result, interaction behaviors of traffic streams between main and minor road frequently occur at priority un-signalized intersections. In this way, vehicles' speed decreases and vehicles dwelling delay time increases as well. Moreover, potential traffic accidents may happen at any time. Since its importance, many attentions have been paid to evaluating the vehicles' interactions between minor and main road. Some research has been done about the speed, acceleration, deceleration, capacity, delay, headway, traffic conflicts, and travel time regarded as the indicators to evaluate the performance of vehicle operation, especially the interactions occur among vehicles. However, passing time of vehicles from the minor road crossing the intersection as the indicator investigating the operation performance of left turning is not yet to be developed.

The primary objective of this study is to develop appropriate quantitative approaches to analyze the interactions of traffic stream between main and minor road based on the passing time regarded as the indicator at un-signalized intersection. More specifically, this study includes two tasks: (1) to compare the changes of passing time 1 and 2 and the whole passing time by applying statistical methods to field data and (2) to make a quantitative analysis of the interaction influence on the passing time and find out how the influential variables act on the passing time. It is hoped that the results can help to enhance the intersection safety and improve the intersection planning and designing.

The paper first reviews the previous literature. The next section explains the typical intersection in Nanjing and describes the survey method, as well as descriptions of videotape coding data. In the third section, statistical methods and

quantitative analysis model are related to evaluating indicators given. Then, statistical results and model validity are discussed. The conclusions and suggestions for future research are summarized in the last section of this paper.

Literature Review

In the past decade, a great amount of appearance indicators has been concerned to evaluating the vehicles' interactions and operation efficiency at un-signalized intersections.

Driver decisions in these situations and the classic procedure for the determination of capacity were based on the calculation of the distribution of critical gaps in mainstream and follow-up times for the vehicles from the minor road at priority un-signalized intersections [2]. And a number of variations on the basic gap-acceptance models have been proposed to estimate delays and capacities at un-signalized intersections [3].

However, the additive conflict flow (ACF) method was developed for more detailed geometric and traffic conditions at two-way stop-controlled (TWSC) intersections to overcome a few drawbacks brought by such gap-acceptance method by Brilon and Wu [4]. Due to the gap-acceptance model did not take the behavior of aggressive or polite driver into consideration. The situation was worsened by heterogeneous traffic stream interactions if priorities are not entirely respected, under a mix of motorized and non-motorized modes [5].

Brilon and Thorsten [6] provided a modified method to determine capacities of TWSC intersections in light of Wu's methodology. Li explored capacity models for estimating capacities of un-signalized intersections under mixed traffic conditions based on the conflict technique [7]. Moreover, an improved model was presented for estimating capacities of vehicular movements based on the conflict technique at multilane AWSC intersections under non-saturated traffic conditions [8].

Estelle presented a dynamic macroscopic traffic flow model for un-signalized intersections that able to predict accurate average vehicle delay and maximum queue length estimates compared to theoretical and empirical data, even under free-flow conditions [9].

Instead, vehicles forcely through with this stream interaction resulting in many conflicts. Traffic conflicts were used as a supplement to estimate the traffic accident potential at intersections or other location and may be considered as a measure of risk [10]. Sayed used the application of the traffic conflict technique for analyzing safety at un-signalized intersections [11]. Kirolos presented multiple approaches to the analysis of crash injury severity at three- and four-legged un-signalized intersections in the state of Florida from 2003 to 2006 [12].

Wang evaluated the normal accelerating behavior of passenger vehicles starting from rest at AWSC intersections by using GPS, verifies the influence of speed limits on acceleration rates, and develops two new polynomial models for driver acceleration behavior [13].

Tiwari applied the survival analysis with nonparametric Kaplan–Meier method to study pedestrians' waiting time at signalized intersections in India [14]. Yang presented a hazard-based duration approach to investigate riders' waiting times, violation hazards, associated risk factors, and their differences between cyclists and electric bike riders at signalized intersections [15].

A number of factors may be contributed to the applicability of models of above indicators. One factor was associated with relatively high main traffic flows, whereby some minor stream vehicles (aggressive drivers) do not wait for gaps in the mainstream. Troutbeck empirically found that mainstream vehicles can be disturbed by aggressive insertion behavior and thus distinguished between the absolute and limited priority modes [16]. Another factor common in some urban areas relates to the low sight distance at intersections, which forces minor road vehicles to take the risk of advancing ahead of the theoretical gap line so as to spot mainstream vehicles, thus forcing mainstream vehicles to give way to avoid collision.

Several important factors affecting crash severity at un-signalized intersections were identified, including the traffic volume on the main and minor approach, and the number of through lanes on the minor approach (surrogate measure for traffic volume) and among the geometric factors [12]. Liu concluded the speed difference between the two vehicles was the most important factor that affected a driver's decision-making. Mostly, the yield occurs when the driver straight-moving from the one side drove significantly slower than the other driver straight-moving across at un-signalized intersections [17].

Pan developed corresponding models about the level of service for safety performance at highway intersection, and the variables are included in the models, such as intersection geometrics, traffic, roadway, and environmental conditions [18]. Wu used the logistic model to study behavior characteristics and associated factors of red light running for two-wheeled riders in China, showing that gender, age, and conformity behavior had significant effects on the cyclist' violation [19]. Rider type, gender, waiting for position, conformity tendency, and crossing traffic flow were identified to have significant effects on riders' waiting times and violation hazards [15].

A review of the literature suggested that some appearance indicators of interactions between main vehicles and minor vehicles have been investigated at an un-signalized intersection, however, the passing time distribution of minor vehicles crossing intersection is often ignored. Thus, new statistical techniques are needed. To fulfill this gap, the study compared the changes of passing time at un-signalized using statistical methods and explored an appropriate and calculable model to evaluate various variables how to affect the passing time.

Data

Test Sites

A cross-sectional observational study was conducted at an un-signalized intersection in Nanjing, China. Four criteria were used to choose the intersection sites as follows. Firstly, the selected sites should be a representative highway un-signalized intersection and not the urban un-signalized intersection. Secondly, there should be a considerably high volume of vehicle traffic and few pedestrian and non-motor volumes. Thirdly, each site around should be tall building that could be convenient to record video. Finally, the selected intersection should be independent and not disturbed by adjacent upstream and downstream intersection. Before the final list of intersections was selected, five signalized intersections were observed and tested.

The typical intersection located in suburban areas in Lishui district of Nanjing was chosen as the observational sites after the pilot. The selected intersection is two-way four-legged on the main road and two-way four-legged on the minor road. They both have segregated barriers and refuge islands which can be used as a safe refuge by pedestrians and riders.

Data Collection

The related data of left-turning behaviors from vehicles were collected at these intersections by placing a video camera. Field observations with video recordings have been widely used to investigate the microscopic behavior, such as analyzing the red light crossing behavior at urban intersections and influences of various types of bus stops on traffic operations of bicycles, vehicles, and buses [20].

Field data were collected on weekdays in during the daytime (i.e., 7:00 a.m.– 6:30 p.m.) under different weather conditions in May and June 2015. And the video data were recorded in peak and off-peak period. There were no traffic wardens employed by the police force.

Videotape Coding

All left-turning vehicles entering the intersection from the Jinlong road to the main road were recorded on video. Right turners were excluded because they were no need to cross the main road at the un-signalized intersection, while left turners were also omitted on account of the limited field of view of the cameras.

Videotape data were recorded in detail by research assistants in the laboratory. Two data files in a predesigned format were made for the site, one for research vehicle from the minor road and the other vehicles from the main road.

Table 1 Definitions of variables coded

Variables	Definition	Descriptions
Minor road		
T_1	Passing time 1	The passing time in decision zone 1
T_2	Passing time 2	The passing time in decision zone 2
T	Passing time	The whole passing time in decision zone 1 and 2
V_i	Vehicle type	0-small 1-middle, big, etc.
Y	Vehicle color	White, black, blue, red, etc.
H_i	Head vehicle	Whether head vehicle entering into decision zone 1 and 2
L_i	Stopping location	0-not stopping , 1–4 as stopping place in decision zone 1
C_i	Whether stopping	Whether the vehicle stopping in decision zone 2
W_i	Weather	1-sunny 2-cloudy 3-rainy
Main road		
V_i	Vehicle type	0-small 1-middle, big, etc.
Y	Vehicle color	White, black, blue, red, etc.
S_{w1}	Speed 1 from the west approach	10 m divide by the gap time of vehicle between section 1 and 2 from the west approach
S_{w2}	Speed 2 from the west approach	10 m divide by the gap time of vehicle between section 3 and 4 from the west approach
S_{e1}	Speed 1 from the east approach	10 m divide by the gap time of vehicle between section 1 and 2 east approach
S_{e2}	Speed 2 from the east approach	10 m divide by the gap time of vehicle between section 3 and 4 east approach
Q_w	Traffic flow rate from the west approach	The vehicle number of driving through the main road from the west approach on decision time in decision zone 1
Q_e	Traffic flow rate from the east approach	The vehicle number of driving through the main road from the east approach on decision time in decision zone 2

The data table for minor vehicles was used to record the variables of passing time 1 and 2 and the whole passing time, vehicles' personal characteristics, movement information, environment factor, etc. In addition, the data table for vehicles on main road was used to record traffic factors related to minor vehicles. Finally, two data files were matched according to the number and type of each vehicle as shown in Table 1.

Generally, vehicle from the minor road entering the intersection (arriving at the stopping line) drivers decide whether to access to the main road directly or wait based on the traffic condition of the west approach on main road, as shown in Fig. 1. It illustrates that drivers need the zone to make a decision on whether to drive into the intersection. Therefore, the zone is recorded as decision zone 1. Vehicles prepare to driving on the opposing main road and adjusted to be straight, drivers decide whether to access to the opposing main road directly or wait based
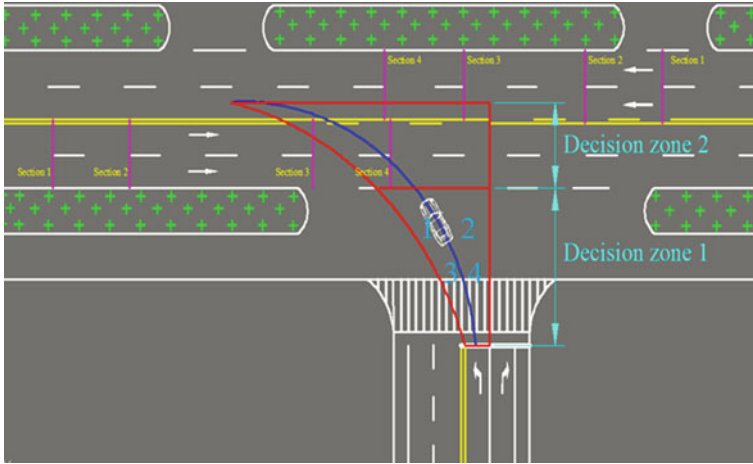
Fig. 1 The diagram at T–J (Tianshengqiao rd.–Jinlong rd.) intersection

on the traffic condition of the east approach on main road. It also illustrates that drivers need the zone to make decisions on whether to pass through the main road. Therefore, the zone is recorded as decision zone 2 as shown in Fig. 1.

The blue curved line represents the normal left-turning trajectory of vehicles from minor road. The red line represents the boundary curve of vehicles left turning, namely normal left turning within the scope of boundary curve and illegal turning without the boundary curve. The green line represents selected section on the main road including two-way approach. The schematic blue numbers are on behalf of rank order of stopping location, following the rule from left to right and forward to backward along the driving direction within specified zone.

Model

Travel time is an essential parameter for evaluating operation efficiency of transportation networks and for assessing the performance of traffic management strategies [20]. In this paper, the survival analysis is the method to study the changes of passing time of minor vehicle's crossing the main road. Survival analysis is a common application in many areas including biomedical, reliability engineering, and social sciences. In the transportation field, survival analysis has been applied to the study of a number of time-related issues including activity duration and vehicle transaction [21, 22].

Hazard-based duration models of survival analysis have an advantage in that it allows the explicit study of describing the duration of a given state and how various factors affect the duration [23]. This is the reason why the duration models are

chosen extensively in biometrics and reliability engineering for decades and also useful tools in the field of transportation.

The process of vehicles left turning can be regarded as a continuous-time state of minor vehicles, which are affected by external factors. Vehicles' passing time of road crossing under the different condition is calculated, and the effects of external factors are quantified. The finding of this paper can build hazard-based duration models to make the quantitative analysis and explain how external factors influence passing time and which factors are more significant.

Survival Analysis and Hazard-Based Duration Model

The variable of interest in duration analysis is the length of time that elapsed from the beginning of an event until its end. In this study, the length of time is the crossing duration of vehicles from the minor road entering the intersection to driving on the opposing main road and adjusted to be straight. The passing time for each rider was taken as the difference between the arrival time when vehicles arrive at the stopping line of intersection and the arrival time when travel to the main road fully.

Let T denote a nonnegative random variable representing travel time of vehicles' road crossing. Let $f(t)$ denote the probability density function of T and the cumulative distribution.

$$F(t) = \Pr(T \le t) = \int_0^t f(u)\mathrm{d}u \tag{1}$$

Let $S(t)$ denote the probability that vehicle travel duration does not end prior to t

$$S(t) = 1 - F(t) = \Pr(T > t) = \int_t^\infty f(u)\mathrm{d}u \tag{2}$$

$S(t)$ is called survival function or survivor probability.

In vehicles' crossing time analysis, the survival function is defined to be the probability that the travel time of road crossing of the vehicle is longer than some specific time, t.

In the survival analysis approach, T can be characterized by a hazard function, $h(t)$. In this paper, it represents the instantaneous probability that a vehicle passes through the main road in an infinitesimally small time period, Δt, after time t, given that the duration has not ended until time t.

$$h(t) = \lim_{\Delta t \to 0} \frac{Pr(t \le T < t + \Delta t | T \ge t)}{\Delta t} = \lim_{\Delta t \to 0} \frac{P(t \le T < t + \Delta t)}{\Delta t \times Pr(T \ge t)} = \frac{f(t)}{S(t)} = \frac{-\mathrm{d}\ln S(t)}{\mathrm{d}t} \tag{3}$$

Note that vehicle travel time is influenced by various factors. The influential factors can be defined as a vector of explanatory variables, $\boldsymbol{x} = (x_1, x_2, x_3, \ldots, x_p)^{\mathrm{T}}$. Then, the proportional hazard form is introduced, which specifies the effects of explanatory variables to be multiplicative on a hazard function

$$h(t) = h_0(t)g(\boldsymbol{x}, \boldsymbol{\beta}) \tag{4}$$

where $h_0(t)$ is called the baseline hazard function and can be interpreted as the hazard function when all covariates are ignored. $g(\cdot)$ is a known function to represent the effects of explanatory variables, and $\boldsymbol{\beta} = (\beta_1, \beta_2, \beta_3, \ldots, \beta_p)$ is a vector of estimable coefficients for $\boldsymbol{x}$. In this paper, a typical specification with $g(\boldsymbol{x}, \boldsymbol{\beta}) = \exp(\boldsymbol{\beta x})$, which was proposed by Cox, is used. The specification is convenient since it guarantees the positivity of the hazard function without placing constraints on the signs of the elements of $\boldsymbol{\beta}$. The Cox hazard model is

$$h(t) = h_0(t)\exp(\boldsymbol{\beta x}) \tag{5}$$

Combining Eqs. 3 and 5, the survival function can be written as

$$S(t) = \exp\left(-\int_0^t h(w)\mathrm{d}w\right) = \left\{\exp\left(-\int_0^t h(w)\mathrm{d}w\right)\right\}^{\exp(\beta x)} = \{\exp(-H_0(t))\}^{\exp(\beta x)} \tag{6}$$

where $H_0(t) = \int_0^t h_0(w)\mathrm{d}w$ represents the baseline of cumulative hazard function; thus, the covariates can be incorporated into the survival function.

Dividing both sides of Eq. (5) by $h_o(t)$ and taking its logarithm, we obtain

$$\log\frac{h_i(t)}{h_0(t)} = \beta x_i = \beta_1 x_{1i} + \beta_2 x_{2i} + \cdots + \beta_p x_{pi} \tag{7}$$

where x_i are covariates for the ith vehicle. The left side of Eq. (7) is a function of hazard ratio (HR), and the right side is a linear function of the covariates and their respective coefficients. The HR can represent the multiple relations between the hazard under the covariate effects and the hazard when all variables are ignored.

Empirical Results

The empirical results are discussed in following two sections. The overall results are presented in first section including model fit statistics and survival probability estimation. The second section presents the effects of covariates.

The Analysis of Passing Time

In this section, the curves of estimated and observed survival probability are presented in Figs. 2 and 3. Figure 2 shows the estimated survival probability by the proposed model and the observed survival probability. Also, the estimated survival probability could be obtained from the Cox model that considers the distribution of passing time under the condition of the average values of all covariates. The nonparametric method is applied to obtain the observed survival probability under the special condition for each driver.

The curve of the estimated distribution monotonously decreases. The median of the distribution is 6.24–5.64 s, indicating that over a half of the observed vehicles can drive into and pass through the tested decision zones 1 and 2 in 6.24–5.64 s, respectively. The 25% quartile of the distribution is 9.12–7.80 s, indicating that about 25% of the observed vehicles cannot drive into and pass through the main road within 9.12–7.80 s, respectively. The estimated results show there are some differences though the general shape is the same between the two results in figures. Specifically, compared to the estimated results, the observed survival probability is smaller until about 9.12–7.80 s, larger thereafter.

The difference existed between observed and estimated survival probability for T_1 is more than T_2, which illustrates that the distribution of T_1 is more complex than T_2 likely because of T_1 influenced by more variables. And the longest survival T_1 in decision zone 1 is above 40 s and far more than longest survival T_2 in decision zone 2. These differences are distributed to be the variable effects, at least partly.

Figure 3 also shows the estimated survival probability by the proposed model and the observed survival probability for the whole passing time. The median of the distribution is 12.36 s, indicating that over a half of the observed vehicles can pass through the decision zones 1 and 2 in 12.36 s. The 25% quartile of the distribution is 17.96 s, indicating that about 25% of the observed vehicles cannot pass through the whole decision zone in 17.96 s. There are also some differences between the two results; specifically, compared to the estimated results, the observed survival
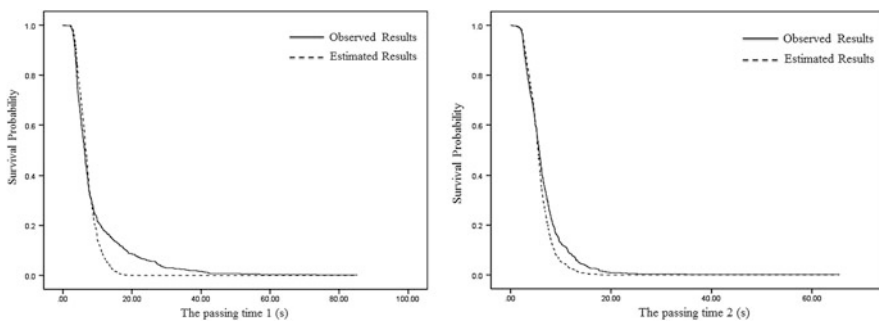
Fig. 2 Observed and estimated survival probability versus the passing time 1 and 2

Fig. 3 Observed and estimated survival probability versus the whole passing time
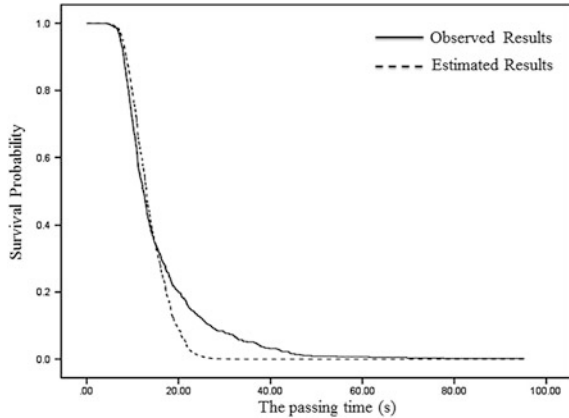
probability is smaller until about 17.96 s, larger thereafter. This difference is partly expected to be the variable effects.

Thus, from the Figs. 2 and 3, the observed results indicate the passing time under the specific condition for individual sample, while the estimated results indicate T_1 and T_2 as well as the whole passing time under the average condition for all the samples. The estimated survival probability reflects the characteristics of the passing time which has an average value for each variable, and any change of the variables could influence the estimated results. Thus, the effects of variables will be introduced in the next subsection.

Effects of Explanatory Variables

Firstly, the statistical test of the estimated model clearly indicates the overall goodness of fit. The method of the forward Wald regression used for model estimation is 5805.673, which is greater than the chi-square value with 7 degrees of freedom at any reasonable level of significance. Considering the interaction of explanatory variables, five, four, and seven variables are selected, and the final estimation of the duration model for the passing time 1 and 2 and whole passing time are shown in Table 2. On the other hand, it can be observed that all of the variables included are statistically significant at the 0.05 level of significance. The variables were selected by the forward selection method. It should be noted that a positive coefficient on a variable implies that the variable increases the passing hazard. With regard to the magnitude of the variable effects, when variable changes by one unit, the hazard would change by $[\exp(\beta) - 1] \times 100\%$.

Table 2 Model estimation of the passing time 1 and 2 and the whole passing time

Variable	Coefficient β	Standard error	Wald value	Sig.	Exp (β)	95.0% CI for exp (β)	
						Lower	Upper
Passing time 1							
W_1	0.300	0.134	5.024	0.000	1.349	1.038	1.753
W_2	0.314	0.127	6.128	0.000	1.369	1.068	1.757
L			160.333	0.000			
L_1	−1.258	0.510	6.086	0.000	0.284	0.105	0.772
L_2	−1.641	0.150	19.926	0.000	0.194	0.144	0.260
L_3	−1.658	0.318	27.272	0.000	0.190	0.102	0.355
L_4	−2.008	0.231	75.713	0.002	0.134	0.085	0.211
H	0.271	0.105	6.728	0.005	1.311	1.068	1.609
Q_w	−0.031	0.006	31.956	0.000	0.969	0.959	0.980
S_{w1}	−0.042	0.008	28.704	0.000	0.959	0.944	0.974
Passing time 2							
W_1	−0.273	0.125	4.745	0.000	0.761	0.595	0.973
W_2	−0.628	0.132	22.622	0.000	0.534	0.412	0.691
C	−0.654	0.157	17.424	0.000	0.520	0.383	0.707
Q_e	−0.010	0.004	7.210	0.005	0.990	0.983	0.985
S_{e1}	−0.028	0.006	19.383	0.000	0.973	0.383	0.707
Whole passing time							
W_1	−0.253	0.131	3.693	0.006	0.759	0.623	0.925
W_2	−0.345	0.133	6.758	0.000	0.843	0.802	0.886
L			115.291	0.000			
L_1	−0.791	0.511	2.396	0.042	0.454	0.167	1.234
L_2	−1.074	0.312	11.882	0.001	0.342	0.185	0.629
L_3	−1.283	0.143	80.793	0.000	0.277	0.209	0.367
L_4	−1.799	0.235	58.842	0.0	0.165	0.104	0.262
H	0.289	0.107	7.252	0.007	1.335	1.082	1.647
Q_w	−0.027	0.006	21.315	0.000	0.974	0.963	0.985
S_{w1}	−0.051	0.008	36.532	0.000	0.950	0.934	0.966
Q_e	−0.011	0.004	8.777	0.003	0.989	0.982	0.996
S_{e1}	−0.020	0.007	8.912	0.004	0.981	0.968	0.993
C	−0.452	0.156	8.362	0.004	0.636	0.468	0.864

The Model of Passing Time 1

As shown in Table 2, the variable coefficient of Q_w is −0.042, and the relative hazard is $e^{-0.042} = 0.959$; thus, the negative coefficient indicates that vehicles with larger volume have a lower hazard (lower passing rate 1 and longer passing time 1). With regard to the magnitude of the variable effects, when the variable of Q_w changes by one unit, the hazard would change by $(e^{-0.042} - 1) \times 100\% = -4.1\%$. As traffic

flow rate from the west approach increases on the main road, the interaction and conflict of vehicles between the main road and minor road increases. Therefore, the larger the vehicle volume is, the smaller crossing gap is. Generally, drivers decide whether to drive at the main road directly or wait based on the traffic condition of on the main road, thus leading to the low-passing hazard in passing decision zone 1.

In addition, the variable coefficient of S_{w1} on the main road is −0.031, also showing a negative effect on the passing hazard. Thus, the vehicle speed is larger with the passing hazard lower. Due to drivers mostly decide whether to access to the main road directly based on the speed perception for oncoming vehicles on the main road. So drivers could be careful in driving if the speed of oncoming vehicles from the west approach is fast. With regard to the magnitude of the variable effects, when the S_{w1} changes by one unit, the hazard would change by $(e^{-0.031} - 1) \times 100\% = -3.1\%$. It can be observed that the traffic flow rate and speed have the slightly negative effect on the passing hazard.

The head vehicle shows a positive effect on the passing hazard. When the head vehicle prepared to access to the main road among minor traffic flows, less time is needed to accelerate or start again than vehicles with conformity behavior because of faster response time and better sight distance and therefore finally leads to the increase of passing hazard.

The effect of stopping locations 1, 2, 3, and 4 on the lateral road line indicates that the presence of vehicles stopping is likely to make vehicles a lower passing hazard than not stopping. Generally, the vehicles stop in decision zone 1 because the traffic flow on the main road was too heavy, so the longer passing time of entering into the main road is spend until the decrease of flow on the main road. Due to the vehicles located on stopping locations 3 and 4 can enter intersection more quickly than vehicles on stopping locations 1 and 2 driving into the main road.

The positive sign of the coefficient on the variables W_1 and W_2 indicates that the variable of two kinds of weathers would make the passing hazard larger than the sunny weather. This is because vehicles' driver could be risky driving behavior on the minor road in the rainy weather and cloudy weather arising from the drivers appear the evident boring mood and psychological tension with visibility dropped significantly [23]. Moreover, this situation could be consistent with relevant research conclusions and video observations.

The Model of Passing Time 2

In addition, the variable coefficient of Q_e is −0.010, and the relative hazard is $e^{-0.010} = 0.990$, and thus, the negative sign of coefficient indicates that vehicles with larger volume have a lower hazard (lower passing rate 2 and longer passing time 2). With regard to the magnitude of the variable effects, when the variable Q_e changes by one unit, the hazard would change by $(e^{-0.010} - 1) \times 100\% = -1.0\%$. As traffic flow increases from the east approach on the main road, the interaction and conflict of vehicles between the main and minor road increases. Therefore,

the merging gap becomes smaller with the traffic volume larger. Generally, drivers decide whether to access to the main road directly or wait based on the traffic condition of on the main road, thus leading to the low-passing hazard in decision zone 2.

The coefficient variable S_{e1} on the main road also shows a negative effect on the passing hazard (longer passing time 2). So the larger the vehicle speed is, the lower the passing time is. This is because drivers mostly decide whether to pass through the main road directly based on the speed perception about oncoming vehicles. Thus, drivers could be careful while driving if the speed of oncoming vehicles from the east approach is fast.

The most significant variable of whether stopping indicates a negative effect on the passing hazard that stopping has a lower hazard than not stopping. And vehicles stop in decision zone 2 due to the traffic flow from the east approach on the main road was too heavy, so the longer passing time of vehicles from accessing into the opposing main road and adjusted to be straight is spent.

The negative sign of variable coefficient showed that the variables W_1 and W_2 are likely to make vehicles a lower passing hazard than sunny weather. This is because vehicles' driver could be careful on the minor road in the rainy weather and cloudy weather, or maybe drivers were with better sight distance on the minor road in the sunny weather so that vehicles passes through the main road quickly.

The Model of Whole Passing Time

In the model estimation of the whole passing time, all variables selected into the model are combined with variables from the model estimation of passing 1 and 2. On the other hand, W_1 and W_2 in model of the whole passing time have the same negative effect on passing hazard with model 2 and have the opposite effect on the passing hazard with model 1; moreover, there are differences between the variables W_1 and W_2 in model of passing time 1 and 2.

As shown in Table 2, the variable coefficient of Q_w with the negative effect on traffic flow rate indicates that vehicles have a lower hazard with larger volume (lower passing rate and longer passing time). And the variable coefficient of S_{w1} on main road is negative, which shows a negative effect on the passing hazard. In addition, the negative sign of coefficient on Q_e indicates that vehicles also have a lower hazard with volume increases. The S_{e1} on main road also shows a negative effect on a passing hazard. With regard to the magnitude of the variable effects, when the variables of traffic flow rate and speed change by one unit, the hazards would change by less than 5%. It can be illustrated that the traffic flow and speed have the slight effect on the passing hazard. More conclusions could be drawn: The interaction and conflict for vehicles on the minor road as well as between the minor and main road increases as traffic flow increases from west approach and east approach on the main road. Therefore, the larger the vehicle volume is, the smaller the crossing and merging gap are. Generally, drivers decide whether to driving into

and passing through the main road directly or wait based on the traffic condition on main road, thus leading to the low-passing hazard in decision zone 1 and 2.

The head vehicle also shows a positive effect on the passing hazard. The positive sign of the coefficient on the stopping locations 1, 2, 3, and 4 on the lateral road line indicates the lower passing hazard than not stopping. The negative significant sign indicates that stopping is likely to make vehicles a lower passing hazard than not stopping. Therefore, the reason that leads to the effect in the passing hazard could be the same above explanation in the model of passing time 1 and 2, respectively.

The positive sign of W_1 and W_2 indicates that the rainy weather and cloudy weather would make the passing hazard larger than the sunny weather in decision zone 1 (shorter passing time 1). This is because vehicles' driver could be risky driving behavior on the minor road in the rainy and cloudy weather arising from the drivers exist the evident boring mood and psychological tension with visibility dropped significantly. However, W_1 and W_2 shows the rainy weather and cloudy weather have a lower passing hazard than sunny weather (longer passing time 2). Maybe vehicles' driver could be careful on the minor road in the rainy weather and cloudy weather, or drivers were with better sight distance on the minor road in the sunny weather in decision zone 2. Thus, the coefficient results of W_1 and W_2 in model of the whole passing time have the same negative effect on passing hazard with model 2 and the opposite effect on the passing hazard with model 1, and thus, it means that the variables in model of passing time 2 are the great contribution to the determination of variable coefficient in the model of whole passing time.

Conclusions

The paper proposes appropriate quantitative approaches to the evaluation of operation performance of minor vehicle's left turning using the passing time as the indicator. The studies were conducted in the tested intersection in Nanjing. The whole passing time was divided into two parts in decision zones 1 and 2, respectively, and the effects of variables are examined by the hazard-based duration approach for the influential analysis of the passing time.

The paper provides several important insights into the determinants of the passing time at the un-signalized intersection. First, the results indicate that the passing time 1 and 2 and whole passing time as the indicators, as well as the variables selected into the model, compose the hazard-based duration model of T_1, T_2, and T. Such results can be reflected by the distribution of the passing time affected by any change of an influential variable.

Secondly, the observed and estimated survival probability for T_1 existed more difference than T_2, and the most passing time 1 is more than the passing time 2. Maybe it is because of the passing time can be influenced by mass of variables in decision zone 1, which lead to the model 1 has more complicated distribution of

passing time than model 2. Additionally, the influential variables related to the characteristics of the passing time should be given full consideration in the planning and designing of intersection facilities and policies.

Finally, a hazard-based duration model was proposed to analyze the influential variables related to the passing time. Based on the results of model estimation, the variables H and W show a positive effect on the passing time 1, whereas Q_w, S_{w1}, and L show a negative effect on the passing time 1. Moreover, Q_e, S_{e1}, W, and C show a negative effect on the passing time 2. In addition, all variables from model estimation of T_1 and T_2 are selected into the model of whole passing time, but the variable of W has the opposite effect on passing hazard between the model of T_1 and T.

In terms of the future work, research with more detailed datasets is required. It is necessary to study the relation between delay and the passing time at the un-signalized intersection. Also, other influential factors should be considered, such as intersection type and size and drivers' psychological factor. Findings from this paper may supplement previous research that has provided inspiration for this study. It is hoped that these results may have a better understanding of passing efficiency and help to plan and design proper policies and facilities at the un-signalized intersection.

References

1. Transportation Research Board (TRB). 1997. Review of international practices used to evaluate un-signalized intersections. *Transportation Research Circular,* 468. National Research Council, Washington, D.C.
2. Heidemann, H. 1997. Queuing at un-signalized intersections. *Transportation Research Part B* 31 (3): 239–263.
3. Transportation Research Board (TRB). 1998. Traffic flow theory—A state of the art report. http://www.tfhrc.gov/its/tft/tft.htm National Research Council, Washington, DC.
4. Brilon, W., and N. Wu. 2002. Un-signalized intersections—A third method for analysis. In: *Transportation and Traffic Theory in the 21st Century—Proceedings of the 15th International Symposium on Transportation and Traffic Theory*, 16–18 July 2002 ed. Taylor, M.A.P., 1–19. Adelaide, Oxford: Elsevier Science Ltd.
5. Prasetijo, J. 2007. Capacity and traffic performance of un-signalized intersection under mixed traffic condition. PhD Thesis, Ruhr-University Bochum.
6. Brilon, W., and T. Miltner. 2005. Capacity at intersections without traffic signal. *Transportation Research Record: Journal of the Transportation Research Board* 1920: 32–40.
7. Li, H., W. Deng, and L. Yu. 2008. Capacities and delays at all-way stop-controlled intersections. *Presented at 87th Annual Meeting of the Transportation Research Board*, Washington, D.C.

8. Li, Haiyuan., Zong, Tian, and Wei, Deng. 2011. Capacity of multilane AWSC intersections based on the conflict technique. *Transportation Research Record: Journal of the Transportation Research Board,* 2257: 111–120. Transportation Research Board of the National Academies, Washington, 17 DC.

9. Chevallier, E., and L. Leclercq. 2007. A macroscopic theory for un-signalized intersections. *Transportation Research Part B* 41: 1139–1150.

10. Parker, M.R., and C.V. Zegger. 1998. Traffic conflict techniques for safety and operations: Observer's manual. FHWA-IP-88-027, Federal Highway Administration, McLean, Va.

11. Sayed, T., and S. Zein. 1999. Traffic conflict standards for intersections. *Transportation Planning and Technology* 22: 309–323.

12. Haleem, K., and M. Abdel-Aty. 2010. Examining traffic crash injury severity at un-signalized intersections. *Journal of Safety Research* 41: 347–357.

13. Wang, Jun., Karen K. Dixon., Hainan, Li., and Jennifer, Ogle. 2004. Normal acceleration behavior of starting passenger vehicles at all-way stop-controlled intersections. *Presented at 83th Annual Meeting of the Transportation Research Board*, Washington, DC.

14. Tiwari, G., S. Bangdiwala, A. Saraswat, and S. Gaurav. 2007. Survival analysis: Pedestrian risk exposure at signalized intersections. *Transport Research Part F* 10 (2): 77–89.

15. Yang, X., M. Huan, M. Abdel-Aty, Y. Peng, and Z. Gao. 2015. A hazard-based duration model for analyzing crossing behavior of cyclists and electric bike riders at signalized intersections. *Accident Analysis and Prevention* 74: 33–41.

16. Troutbeck, R.J., and S. Kako. 1999. Limited priority merge at un-signalized intersections. *Transportation Research Part A* 33 (3–4): 291–304.

17. Liu, M., G. Lu, Y. Wang, and Z. Zhang. 2014. Analyzing drivers' crossing decisions at un-signalized intersections in China. *Transportation Research Part F* 24: 244–255.

18. Lu, Jian., Fuquan, Pan., Qiaojun, Xiang., and Guoqiang, Zhang. 2007. Level of safety service for safety performance evaluation of highway intersections. *Presented at 87th Annual Meeting of the Transportation Research Board*, Washington, DC.

19. Wu, C.X., L. Yao, and K. Zhang. 2012. The red-light running behavior of electric bike riders and cyclists at urban intersections in China: An observational study. *Accident Analysis and Prevention* 49 (11): 186–192.

20. Vanajakshi, L.D., B.M. Williams, and L.R. Rilett. 2009. Improved flow-based travel time estimation method from point detector data for freeways. *Journal of Transportation Engineering* 135 (1): 26–36.

21. Habib, K.M. 2012. Modeling commuting mode choice jointly with work start time and work duration. *Transportation Research Part A* 46 (1): 33–47.

22. Nam, D., and F. Mannering. 2000. An exploratory hazard-based analysis of highway incident. *Transportation Research Part A* 34 (2): 85–102.

23. Zhu, Z., J. Rong, and W. Zhou. 2010. The driving behavior study under bad weather conditions of. *Journal of Wuhan University of Technology (Transportation Science and Engineering Edition)* 12: 1040–1043.

The Detection and Precision Analysis of Bright Color Vehicles Based on Remote Sensing Image

Dudu Guo and Jing Zhang

Abstract The acquisition methods of traditional traffic information have many limitations. Getting a wide range of road traffic information from the remote sensing image becomes a hot research topic of the information acquisition technology in intelligent transportation system. First, this chapter enhanced the remote sensing image to improve the quality of the images; second, it is proceeded with multiscale image segmentation; then used the adjacent classification to classify the image carried out on the segmentation, selected vehicle samples and not selected vehicle samples, and trained classifier to realize image classification; and finally, it evaluated the detection results.

Keywords The remote sensing image · Bright color vehicle extraction · Classification · Precision

Introduction

The existing traffic sensors to collect traffic information exist many limitations [1], such as the high cost of equipment installation and maintenance, low data accuracy, small coverage, and incompatibility with each sensor data evaluation system , especially in terms of access to a wide range of traffic parameters, where the traffic parameters is limited. With the emergence of commercial high-resolution satellite, a wide range of road traffic information from remote sensing image has been put into the possible of the information acquisition technology in intelligent transportation system [2].

D. Guo (✉) · J. Zhang
College of Mechanical Engineering, Xinjiang University, Urumqi, China
e-mail: guodudu1122@126.com

D. Guo
Traffic College, Wuhan University of Technology, Wuhan, China

The Extraction Method of Bright Color Vehicles from Remote Sensing Image

We can distinguish between different objects quickly because of the different characteristics of the object that the human eye can distinguish the target from the background. In remote sensing image target recognition technology, the choice of what kind of characteristics to describe the content of the image of target is the key to get more accurate results.

Structurally, the inhomogeneous vehicle target compared with homogeneous road, spectral change frequency is higher and the texture information is more abundant [3]. In dense fleet vehicle, arrangement has certain regularity. Its texture structure is more obvious. Between vehicle and road, shadows and other vehicles and road environment also exists certain spatial relations: such as vehicle traffic direction and roads in the same direction, or both in a certain Angle range; In the same direction of vehicle traffic lane; Having a certain distance between vehicles and vehicles; Vehicle with its own shadow always going together. So the selected characteristics have the direct impact on vehicle detection results. The key is to choose the accurate reflection of the characteristics of the vehicle information to improve the classification accuracy.

The Image Enhancement

The enhancement of remote sensing image can improve image visual effect, enhance image differences or characteristics of target object, and highlight the target feature from the environmental background information processing [4]. When analyzing remote sensing images, in order to make the analysis easy to identify the image content, it should be carried out in accordance with the analysis of the purpose of image data processing to improve the image of interpretation.

There are two main purpose of image enhancement: One is to improve the image visual effect and the clarity of the image; the second is to transform the image into a more suitable for human or machine analyzed in the form of processing to get more useful information from the image. The point calculation formula is as follows:

$$g(x, y) = I[f(x, y)] \tag{1}$$

where I is the linear or nonlinear transformation of (x, y), but not the location.

The point calculation is to put the gray value of each original image pixels into a new gray value of the output image through the transformation. The position of each original image pixels in the image does not change.

Image Segmentation

The basic idea of multiscale image segmentation is: starting from a single pixel, respectively merging with its neighborhood and calculation to reduce the heterogeneity of the final result [5]. After the end of the first merger round, regarding the generated objects as the basic unit, it continues to work with its neighbors' combined objects and calculated separately. This process will continue the user specified scale has no longer any object merge. In the practical application, the close value of the heterogeneity degree needs to be defined by setting the integral scale parameter. Within the closing value, it will merge, otherwise not be merged. So it needs to choose the appropriate parameters according to the classification of the different target, in order to get reasonable segmentation result.

Multiresolution segmentation uses different scales to form network hierarchy. Each partition uses low layer of image object as a raw material, and the material merges in the new division. At the same time, it also follows the higher layer of object boundary limit. The network structure is a topological relation. For example, the boundary of the parent object determines the boundary of the object, and the size of the parent object region is determined by the sum of child objects. Each layer is composed of its immediate child objects. On the next top, child objects emerge into large objects. The merger will be limited by the boundary of the parent object. If the parent object is different, the adjacent objects cannot be merged. Starting from a single pixel, adjacent object is calculated, respectively. If the two objects after the merger of adjacent heterogeneity index are less than the given threshold, they will be merged, otherwise not be merged.

The Image Classification

Image classification is the main purpose of the satellite images to detect vehicles. The vehicle detection is essentially from the image recognition, and to get the related information. So the image classification is the key technology of vehicle detection. Vehicle detection task from various kinds of features can be seen as the vehicle and the no-vehicle classification problems.

The adjacent object classification looks for the latest samples in the feature space object [6]. If an image object of the nearest sample objects belongs to A class, then the object will be divided into class A. In actual operation, by class A membership function. If the image object in the feature space near a sample belongs to the class A, the membership degree that belongs to A class is higher.

The Extraction of Bright Color Vehicles

Scale parameter is an abstract term which can decide the final image maximum heterogeneous degree of objects. With the given scale parameter, heterogeneous data generated object segmentation is much smaller than in homogeneous data. Changing the size of the scale parameter can obtain different sizes of image objects. Figure 1a shows the original pictures of Geo Eye multispectral images with a resolution of 0.5 m and 160 * 218 pixels. The segmentation result with different scales is shown in Fig. 1b–d.

The images after multiscale segmentation then need to be classified by the standard adjacent object classification. Common classification features of remote sensing image are the shape feature and texture feature. Multiscale segmentation is to establish the relationship between each object and provides rich image object characteristics. The subclassified correction of remote sensing image is shown in Fig. 2.

(a)	(b)	(c)	(d)
The original image	Segmentation image	Segmentation image	Segmentation image
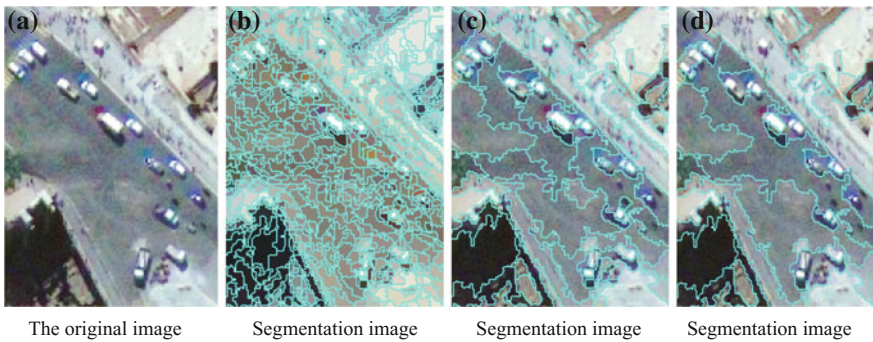
Fig. 1 Multiscale segmentation image of Geo Eye satellite image

Fig. 2 Classification of remote sensing image of Dallas Airport
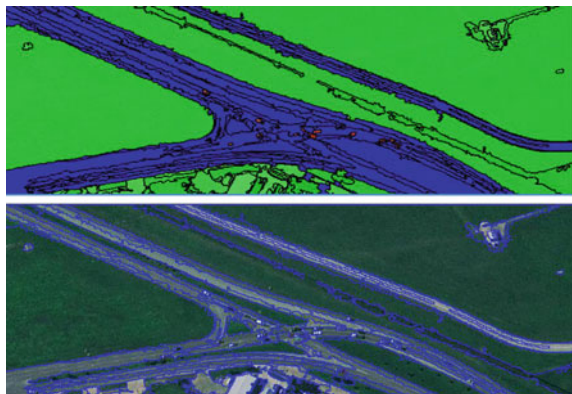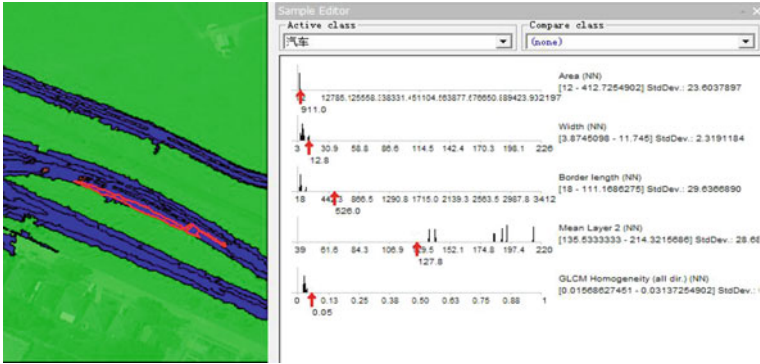
Fig. 3 Gray value of remote sensing image classification of Dallas Airport

(a)

(b)
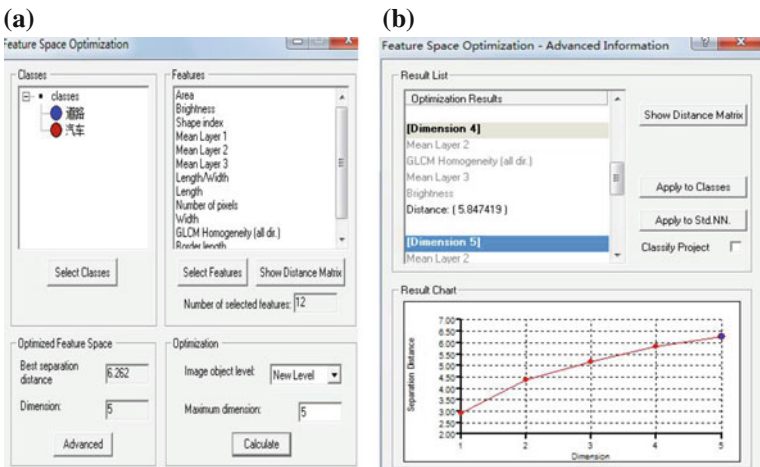
the selected features   the image separation results of Dallas airport

Fig. 4 Vehicle detection result

In classification, sample editor red pointer will indicate the standard in the adjacent classifier characteristics of the sample under the corresponding gray value, including mean value and variance of image objects, gray, brightness, and other characteristics. By checking the same classification characteristic value of each child objects, the feature threshold of the same classification can be obtained. The classification characteristics of gray value of the remote sensing image comparison are shown in Fig. 3.

Vehicle test result is shown in Fig. 4.
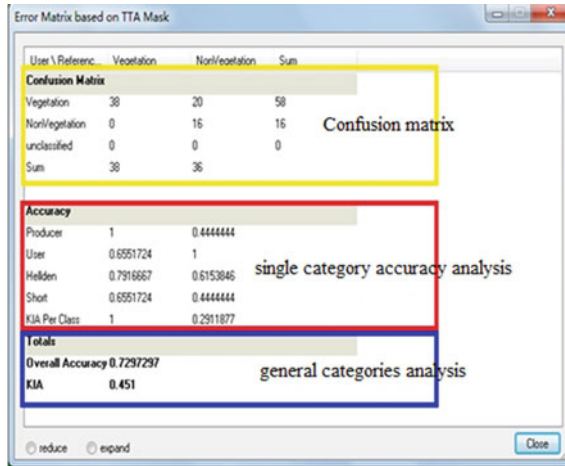
Fig. 5 Accuracy analysis indicators

Fig. 6 Accuracy evaluation results of Dallas Airport

Accuracy Evaluation

Accuracy evaluation includes three parts: confusion matrix, the precision of the single-category analysis results, and general categories of precision analysis results. Confusion matrix can be used to understand each category of the total number of samples and wrong number of sample points and leakage; single-category accuracy analysis includes five indicators; general category analysis includes two indicators. They are shown in Fig. 5.

The accuracy evaluation results of the above examples are shown in Fig. 6.

For the single-category precision analysis, the KIA per class of vehicles is 0.6666667, which accounts for the classification results with good quality. Car user accuracy is 100%. The general category precision is 0.9.

Conclusion

The test result of bright color vehicle detection rate is high, and it is because the spectrum contrast of bright color vehicle is big relative to the road surface, and geometric feature is obvious, which is very beneficial to the vehicle detection. But the test results are greatly influenced by shadows, which is the problem existing in the research.

References

1. Ming, Dongping, Jiancheng Luo, and Zhanfeng Shen. 2005. High resolution remote sensing image information extraction and target recognition technology research. *Journal of Surveying and Mapping Science* 30 (3).
2. Wang, Weichao, and Weibao Zou. 2013. High resolution remote sensing image information extraction method review. *Journal of Surveying and Mapping* 13 (4) (Beijing).
3. Liang, Yanping, and Mingliang Suo. 2011. The new direction and progress in the study of sport vehicle detection. *Journal of Wuhan University of Technology* 35 (4) (Transportation Science and Engineering Edition).
4. Luo, Bo. 2013. *High resolution remote sensing image segmentation method research*. University of Electronic Science and Technology.
5. Qi, Sun. 2010. *Based on high resolution satellite images of traffic flow parameters extraction research*. Beijing Traffic University.
6. Hao, Huang. 2005. *Based on high resolution satellite imagery of urban vegetation information extraction and analysis research*. Hehai University.

Segment Division Analysis of the Low-Grade Roads Traffic Flow Access in Suburban Area

Ning Zhong, Yan-kai Zong, Yong-feng Ma and Jian Lu

Abstract With the economic development, accelerated urbanization process, constant urban expansion, and stronger land exploitation, the traffic problems of suburban area become increasingly serious. Compared with the access entry of urban roads, the problem of single access entry of low-grade roads in suburban area is not outstanding, but the mutual impact of the access entry clusters has a huge influence on the mainline. In order to study the management methods of access entries in distinct attributes of low-grade roads in suburban area intuitively, segment division shall be carried out for low-grade roads in suburban area. Through the analysis of specific sections, targeted solutions can be proposed, which may reduce the macroscopic countermeasures of unclear directivity and reduce the deviation of research result due to the subjectivity. In this paper, double-layer division mode is taken for the research, and the different section types obtained from the research can be the basis of subsequent studies.

Keywords Suburban area · Low-grade road · Access · Section

Introduction

It is found out that roads of suburban area are mainly low-grade roads whose technical grade is level three or below, while sections refer to a road with one or several characteristics, for instance, the same access entry, speed discrete section, and concentrated conflicting points. According to studies on low-grade roads in suburban area carried out by numerous scholars. Due to the impact of various factors, such as different land types, road conditions, traffic characteristics, traffic safety, and access entry type, the low-grade roads in suburban area may have

N. Zhong · Y. Zong · Y. Ma · J. Lu (✉)
School of Transportation, Southeast University, Nanjing 210096, Jiangsu, China
e-mail: lujian_1972@seu.edu.cn

correspondingly different access management methods. Types of different sections can be obtained through the segment division, and targeted analysis is carried out. Besides, specific solutions are proposed, for reducing the errors of research result caused by the subjectivity.

The segment division of roads, according to the section characteristics, is the foundation of generating access management plans and realizing the coordinated management. Meanwhile, the appropriateness of the segment division will directly determine the scientificity of the access management plan.

Characteristics of Road Section

Currently, there have not been specialized studies on the segment division of roads at home or abroad, and road section is mainly studied in a specific form, for instance, accident-prone section, dangerous section, and demonstrative section, but there are few studies on the division of the entire road.

The road section of this research mainly refers to roads with various characteristics, for instance, similar basic road conditions, surrounding land development nature and strength, similar speed discrete degree, traffic accidents, and traffic conflicts. Furthermore, the section contains two or more access entries. By defining the road section and combining the traffic characteristics, the road section has the following characteristics:

- The length of road section is unfixed;
- The land use along the road section is similar;
- The fundamental road condition is similar;
- The traffic flow is of similar characteristics;
- The traffic safety is similar.

Segment Division of Roads

General Rules of Classification

Based on the practical conditions of low-grade roads in suburban area, land utilization and development forms, basic road conditions, traffic characteristics, and traffic safety, taking advantage of the double-grade division mode, qualitative division for the section type is carried out according to the land utilization development and basic road conditions. Later, quantitative division is carried out by indexes such as the speed and traffic conflict, and different section types are formed eventually.

Qualitative Divisions

Study on Division Methods

Qualitative division research covers two aspects, which are separately land utilization development and basic road conditions. Process statistical analysis by quantized data. Qualitative division is mainly through consulting document literatures and combining the field research.

Impact on the road traffic of the surrounding land usage of access entry and basic road conditions is analyzed by studying domestic and foreign studies on the roads of suburban area, and sections with similar impact on the traffic mainline are divided into the same type.

Specific Division

In accordance with the *Road Engineering Technical Standard* (JTGB01-2014), it can be concluded that if low-grade roads in suburban area are of the same grades, the basic road conditions would be similar.

It is found out that lands on both sides of low-grade roads in suburban area mainly consist of residential land, commercial land, industrial land, and agricultural land according to the field research of roads in domestic suburban area and *Urban Land Classification and Planned Construction Land Standard*. Villages are formed in commercial and residential lands, while agricultural lands mainly cover farmland, forests and fish ponds, etc. If there are villages along the low-grade roads in suburban area, this section is named as village-through section; correspondingly, sections dominated by agricultural land are named as the common section. The section between the general section and village-through section is the transition section, named village-impact section, and it is similar to the influence area of intersection. The village-impact section and village-through village can be separated by place of residence.

Therefore, as shown in Fig. 1, roads can be divided into common section, village-impact section and village-through section qualitatively. On this basis, each type of section will be refined according to the land usage and development intensity, for instance, the village section includes different types of land, such as commercial district and residential area; as for the impact on traffic safety and expedite, it is gradually reinforced from the common section to the village-through section.

Fig. 1 Schematic of road type
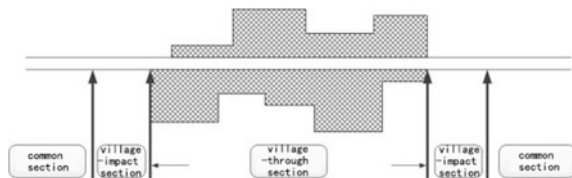
Quantitative Divisions

Study on Division Methods

The probability of traffic accident is not in simple increase with the speed, but the "U-shape" curvilinear relationship [1] ,as shown in Fig. 2. When the speed approaches the mean traffic flow, the rate of traffic accident would reach the minimum, while when the speed deviates from the mean traffic flow, the probability of traffic accident would increase correspondingly with the increase of deviation.

Therefore, the speed would impact the road traffic safety, and there is a close relation between the speed discreteness and traffic conflict, which may impact traffic safety as well. In this paper, the speed discreteness degree and traffic conflict are selected as the index of traffic safety.

(1) Speed factors

The difference of vehicle speed and mean speed of traffic flow is reflected by the speed discreteness of traffic flow. In this research, the speed discreteness is specifically reflected by the speed differences in the same location at different moments. In the current studies, the speed discreteness is expressed in speed standard deviation (SD):

$$\text{SD} = \sigma = \sqrt{\frac{\sum_{i=1}^{n} (v_i - \bar{v})^2}{n - 1}} \tag{1}$$

In which, v_i is the speed of the ith vehicle, and $\bar{v}$ is the average speed of all vehicles.

The mean speed of each section has a huge impact on the speed discreteness. In this paper, the evolving form of the SD, "speed standard variable (CV)," is selected for eliminating the impact of speed on the speed discreteness:

$$\text{CV} = \sigma / \bar{v} \tag{2}$$

Fig. 2 Diagram of speed and accident rate
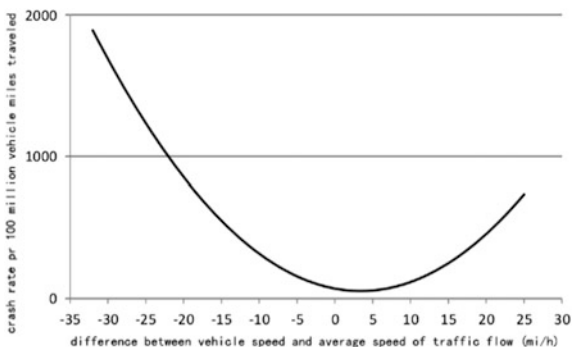
The speed standard variable can be applied for describing the discreteness, and it is the standardization of standard deviation.

(2) Conflict factors

The degree of traffic flow conflict is studied by conflicts, and the traffic volume and conflicts can be obtained according to the investigation of all locations. Since different locations have distinct traffic volume, it may have a huge impact on the research result. In this study, conflict ratio (TS) is employed as the influential factor in the study on conflict degree [2]:

$$TS = T_i/V_i \tag{3}$$

In which, T_i is the number of conflict, and V_i is the flow of point i. The conflict ratio can eliminate the impact of different traffic volumes on the research result, which makes the research result applicable extensively.

The speed discreteness and degree of conflict of all locations in the section will present different characteristics, which can be divided into different types through analysis. Besides, the corresponding section of speed discreteness and degree of conflict of different types is also a distinct section.

The classification standard of speed standard deviation CV is obtained through the statistical analysis of speed data of all measuring points; the speed discreteness of current road is divided according to the classification standard, and different types correspond to distinct road sections, namely the road segment division. Different types of sections would be obtained according to the distinct types of access entry within the section.

Specific Division

(1) Segment division

In this paper, the sample section selected is a two-lane road, without separation between motor and non-motor and central partition. According to the investigation, the mean speed of the sample section is 60 km/h, the response time of the driver is 2–3 s internationally. Therefore, within 50 m, the speed change is relatively small. Consequently, statistical points are set every 50 m for data collection, and such data as the speed, flow, and conflicts within 50 m of all section locations are counted. If the location is near the access entry, move the point to the access entry and related data of the access entry shall be counted, as shown in Fig. 3.

Fig. 3 Schematic of sample section data collection
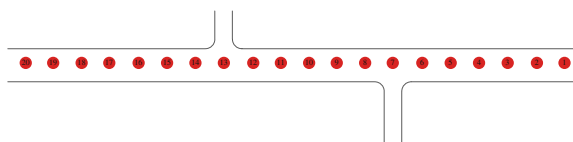
Fig. 4 Diagram of speed
standard deviation variable
coefficient and traffic conflict
ratio
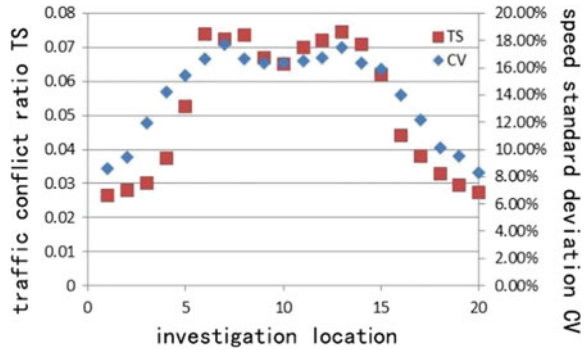
In each location, the speed of 100 vehicles shall be counted. The conflicts in the morning and evening peak between two locations shall be counted, once in 15 min, total 50 groups, as shown in Fig. 4:

It is illustrated in Fig. 4 that:

1. There is a good correlation between the traffic conflict ratio and speed standard deviation CV, and two factors impacting the quantitative division of the section share similar division effect. In the following study of segment division, the speed standard deviation CV is selected as the main factor for analysis.
2. When vehicles approach the access entry, the standard deviation CV increases and reaches the maximum near the access entry. It maintains a high level between two access entries, suggesting the mutual impact of two access entries, and these two entries should be divided into the same section.
3. The curve is relatively stable when the position of vehicles is at location 1, 2, 3, 17, 18, 19, and 20, and it increases or decreases rapidly from location 4 to access entry and location 16 to access entry in a relatively steep way.
4. The corresponding standard deviation CV of location 4 and 16 is 14.26 and 14.01%, respectively; therefore, the classification standard shall be between 14 and 15%.

Therefore, the classification standard of CV ranges between 14 and 15%, and if it is not lower than the classification standard between two access entries, it can be considered that the two access entries can be divided into the same section, otherwise the two are in different sections.

(2) Influence scope of the access entry

The current research results have failed to form systematic studies on the influence region of the access entry. In this paper, the influence region of the mainline of access entry is determined by referring to the influence area of intersection.

The intersection mainly includes physical area and functional area of the intersection. The former refers to the region enclosed by the stop line of all accesses, which allows vehicles to be in and out of the intersection; the latter is the

Fig. 5 Schematic of intersections
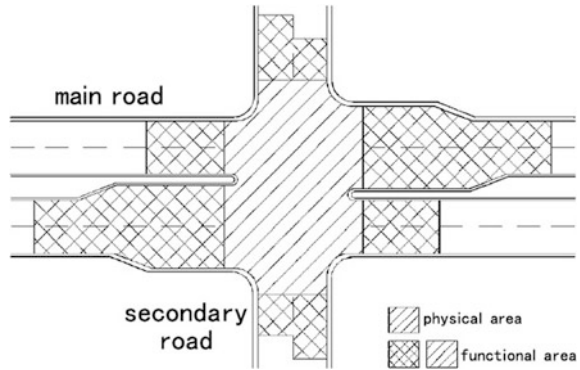
Table 1 Design length of the functional area in intersections

Design speed (km/h)	Length of upstream functional zone (m)	Length of downstream functional zone (m)
100	280	180
80	210	125
60	155	85
40	110	50

extended section allowing vehicles to drive in or out of the intersection, and it is mainly intended for the acceleration, deceleration, change of routes, waiting in line, etc. as shown in Fig. 5:

The determination of the influence scope of the intersection shall be the determination of the length of the functional zone. Since it is relatively mature in domestic and foreign research achievements, it will not be studied separately. According to different speed limits, the length of functional area is also distinct [1] (Table 1):

The speed limit of the sample section is 60 km/h, and it can be determined that the impact length of access entry to the mainline is 155 m. In the segment division, the corresponding distance of CV ranging between 14 and 15% is about 150 m. In order to guarantee the traffic safety of roads at the access entry, the longest distance is selected as the influence length. Combining the setting of CV location, the influence distance is confirmed as 150 m, and the corresponding CV is 15%, which is similar to the research result of Del Castillo and Benitez [3].

In conclusion, the classification standard of speed standard deviation CV is finally defined as 15%.

(3) Classification significance test

In order to guarantee the accuracy of research result, significance test is generated for the division of the sample section. Since CV classification standard is 15%, it can be learnt from Fig. 4 that locations 4 and 5 are in different sections. If

Table 2 Conflict statistical analysis table

Conflict rate Type	Number of samples	Mean value	Standard deviation
Location 4	50	0.03747	0.00512
Location 5	50	0.05256	0.00652

Table 3 SPSS variance analysis result

	Sum of squares	df	Mean square	F	Significance
Inter-group	0.006	1	0.006	162.436	0.000
Intra-group	0.003	98	0.000		
Total	0.009	99			

the difference between the two is proved as significant, it means that the classification is valid. The test data adopts the conflict rate data, and the treatment result is shown in Table 2:

In order to verify if there are differences between two locations, analysis of variance(ANOVA) shall be processed, and it is realized with the assistance of SPSS software. The result is shown in Table 3:

It can be learnt from the table that the significance is smaller than 0.05, suggesting that when the confidence interval is 95%, it is considered that there is significant difference between the two locations, which belong to different sections, and it also suggests that the classification standard of CV about 15% is reasonable.

(4) Section type

There are various forms of access entries on both sides of the low-grade road in suburban area, and different types of access entry have a distinct impact on the speed and conflict degree of vehicles in the main road. Therefore, based on the segment division, different types of access entry contained in each section determine the type of section, and the division of section type requires firstly to divide the type of access entry.

According to the standard of land use along the roads in *Urban Land Classification and Planned Construction Land Standard* (GB50137-2011), the specific classification of the access entry in American Entrance and Exit Management (7 categories and 5 types), as well as the study achievements of exit and entrance classification [4], combining the field survey of low-grade road in the suburban area, the access entry is divided into four types, namely the access entry for large-scale business and community, access entry for villages and branches of main roads, access entry for agricultural and forestry land, and access entry for small shops and homes along the street (Table 4).

The section type is divided by combining the category of access entry. There are mainly the following 14 types of road section types (Table 5):

Table 4 Division of access entry types

Category of access entry	Type of access entry	Introduction
Category I	Access entry for large-scale business and community	Vehicles on the access road are mainly dominated by passenger cars, vans, and electric vehicles, with a huge flow. It has a huge impact on the traffic of the main road, and there are many traffic conflicts
Category II	Access entry for villages and branches of main roads	Vehicles on the access road are dominated by agricultural vehicles and non-motor vehicles. There are few vehicles, however, the speed is high. Parts of sight distances are poor, which can give rise to the traffic accidents easily
Category III	Access entry for agricultural and forestry land	Vehicles on the access road are dominated by agricultural vehicles and non-motor vehicles. It is mainly dominated by agricultural production, and vehicles of the main road usually accelerate to pass
Category IV	Access entry for small shops and homes along the street	Types of cars in the access entry are complicated, and there are numerous access entries, with dense pedestrians and traffic, which generate a huge impact on the traffic of the main road

Table 5 Statistical form of road section type

No.	Type of access entry within the section	Type of access entry within the section
1	Type I	Access entry for large-scale business and community
2	Type II	Access entry for villages and branches of main roads
3	Type III	Access entry for agricultural and forestry land
4	Type IV	Access entry for small shops and homes along the street
5	Type I and II	Access entry for large-scale business and community Access entry for villages and branches of main roads
6	Type I and III	Access entry for large-scale business and community Access entry for agricultural and forestry land
7	Type I and IV	Access entry for large-scale business and community Access entry for small shops and homes along the street
8	Type II and III	Access entry for villages and branches of main roads Access entry for agricultural and forestry land
9	Type II and IV	Access entry for villages and branches of main roads Access entry for small shops and homes along the street
10	Type III and IV	Access entry for agricultural and forestry land Access entry for small shops and homes along the street
11	Type I, II, and III	Access entry for large-scale business and community Access entry for villages and branches of main roads Access entry for agricultural and forestry land

(continued)

Table 5 (continued)

No.	Type of access entry within the section	Type of access entry within the section
12	Type I, III, and IV	Access entry for large-scale business and community Access entry for agricultural and forestry land Access entry for small shops and homes along the street
13	Type II, III, and IV	Access entry for villages and branches of main roads Access entry for agricultural and forestry land Access entry for small shops and homes along the street
14	Type I, II, III, and IV	Access entry for large-scale business and community Access entry for villages and branches of main roads Access entry for agricultural and forestry land Access entry for small shops and homes along the street

In road sections, there are generally two or more access entries, which mainly in one or two types. Cases in which a section contains three or four types of access entry are few.

Case Study

Zhejiang Suichang suburban area Sanji line (X603) is about 109.803 km length in total, among which, the design speed of 0 and 13.5 k is 60 km/h. The road is dominated by bidirectional two-lane section and single carriageway, while some are bidirectional four-lane section and two-carriageway. In this paper, Yaobu Floor Factory section is selected as a case for analysis.

Traffic Survey

It is about 600 m length in total, and 12 data acquisition points are set. As for land use types, in the west is the agricultural land, while in the east is the commercial land, distributing many factories and enterprises. The road is bidirectional four-lane section and two-carriageway, with central partition contains seven openings. Among which, No. 1, 2, 3, 5 6 accesses are similar to Type I access entries, while No. 4 and 6 are type II access entries, as shown in Figs. 6 and 7:

Sector Divisions

Results obtained through the analysis of the data acquired are showed in Table 6:

Fig. 6 Schematic of the current roads of Yaobu Floor Factory section

Fig. 7 Schematic of the data acquisition points of Yaobu Floor Factory section
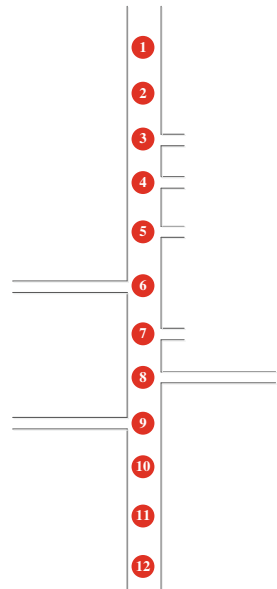
It can be concluded from the table that the mean speed standard deviation CV is higher than 15% from acquisition point 2–10, suggesting that the two points are in the same section.

Table 6 Data processing result of Yaobu Floor Factory section

No.	Average speed (km/h)	Standard deviation	Speed standard deviation CV (%)
1	62.13	9.25	14.89
2	61.11	9.65	15.79
3	58.43	9.54	16.33
4	56.86	10.08	17.72
5	58.11	10.64	18.31
6	58.08	10.98	18.90
7	58.90	11.43	19.41
8	57.74	11.43	19.79
9	56.68	10.69	18.86
10	57.93	10.61	18.31
11	60.89	8.99	14.76
12	60.46	8.54	14.12

Conclusion

In this paper, detailed studies are conducted for the division of low-grade road in suburban area, and the classification method contains two levels, namely the quantitative and qualitative division.

The qualitative division mainly through the classification standard from such aspects as the land use, development strength, and basic road conditions to generate preliminary division results of the section, which are separately village-through section, village-impact section and common section.

The quantitative division mainly selects CV and TS to analyze the attributes of the section through the study on the sample section, so as to analyze the relationship between CV and TS. Meanwhile, the classification standard of CV is determined as 15% through determining the segment division standard with CV and combining the influence scope of access entries.

In the quantitative division, CV is irrelevant to the road traffic flow, and TS eliminates the impact of traffic flow on the result. Consequently, the road segment division method is applicable for roads with huge traffic flow or small traffic flow theoretically.

14 different types of road sections are finally confirmed by combining the division of the access entry types. The main traffic characteristics of different section types are obtained through analytical study. Distinct types of section contain various quantities or types of access entry, and the traffic problem incurred may also be different. Distinct traffic problems shall be managed with appropriate traffic management technologies. As for the subsequent studies on the access management of low-grade roads in suburban area, the research result of this paper can be the basis and foundation for the formulation of management plans.

References

1. Stuster, Jackand, and Zail Coffman. 1998. *Synthesis of safety research related to speed and speed management*. Publication No. FHWA-RD-98-154, 7.
2. Ma, Yongfeng. 2007. *Study on the reasonable spacing of highway intersections*. Southeast University.
3. Del Castillo, J.M., and F.G. Benitez. 1995. On the functional form of the speed-density relationship part two: Empirical investigation. *Transportation Research B* 29: 391–406.
4. Li, Chao, Rong Jian, et al. 2004. The impact of arterial highway access density on traffic safety. *Collected Papers of International Road Safety Seminar 2004*, 32.
5. Fu, Chengwei, et al. 2010. Review on urban-rural linkage in China. *Progress in Geography* 12.
6. Zhang, Tiejun, Tang Chengcheng, Luo Xun, et al. 2009. Analytical investigation on safety influence of grade highway's accesses in plain area. *Journal of Highway and Traffic Research and Development* 26 (6): 119−122.
7. Xu, Song. 2009. *Research on accesses of first-class highway in suburb*. Chongqing Jiaotong University.
8. Xiaoping, Guang, Qian Yongsheng, and Zhou Bo. 2006. Design research on road traffic safety in suburban area. *Communications Standardization* 2006 (5): 131−135.
9. Conggao, Lv, Yingna Wei. 2007. Analysis and countermeasures of road traffic safety in Chinese suburban area. *Communications Standardization* 12: 180−183.
10. Ma, Xiangjuan. 2008. Traffic problems and planning emphases for suburban area. *Technology and Economy in Areas of Communications* 03: 114–115.
11. Xu, Wangying, Xu Li, Yang Min. 2008. Traffic planning research in suburban area. *Channel Science* 09: 30−31+52.
12. He, Qicheng. 2008. *Research on suburb highway safety*. Chongqing Jiaotong University Master's Thesis, 4.
13. Ma, Junyong. 2009. Traffic design research in suburban area. *Communications Standardization* 15: 124–126.

Study on the Geometric Design Method of Access on the Urban Fringe and Villages Segment

Yue Chen, Da Wu, Yongfeng Ma and Jian Lu

Abstract With the process of urbanization rapid progress and the increase of residents' travel demand and motorization, the access design speed and the key geometric design parameters of the level and cross-sectional plane were recommended according to the typical traffic safety environment and traffic operation features on the urban fringe villages and low-grade road section of China. With the simulation of the geometric improvement on the urban fringe villages and low-grade road section access, the design method and the recommendation of geometric design parameters were verified to be rational and effective.

Keywords Access · Design speed · Safety improvement

Background

With the rapid development of urbanization process and the increase of motorized traffic levels, all kinds of traffic demands grow rapidly and the conflicts among transportation, such as motor vehicles, non-motor vehicles, and pedestrians,

Y. Chen · D. Wu · Y. Ma · J. Lu (✉)
Jiangsu Key Laboratory of Urban ITS, Southeast University,
Dhaka, Bangladesh
e-mail: lujian_1972@seu.edu.cn

Y. Chen
e-mail: chen1993yue@163.com

D. Wu
e-mail: wuda19900610@163.com

Y. Ma
e-mail: mayf@seu.edu.cn

Y. Chen · D. Wu · Y. Ma · J. Lu
Jiangsu Province Collaborative Innovation Center of Modern Urban
Traffic Technologies, Sipailou #2, Nanjing 210096, Jiangsu, China

increase in the majority of village urban fringe and some neighboring towns in China, which lead to the serious situation of road safety. There are several problems in the urban fringe and village roads, such as low design standards, chaos of access management, complexity of traffic composition, inadequate safety facilities, low safety consciousness of the traffic participants, relatively weak investment, and efforts of traffic enforcement [1]. How to effectively improve the level of traffic safety is the difficulty in traffic planning, design, and management [2].

As an effective means to enhance road safety, access management technology is an important part to improve the urban fringe of towns and low-grade sections of road traffic safety environment. Access management technology has carried out extensive research and practical application in Europe and the USA. In 2003, Transportation Research Board introduced the "Access Management Manual," elaborating on the content access management including principles of access management, selection of the access location, spacing of access, geometry design methods of access, traffic organization, approval process of access management, and public participation. This manual provides specific and exemplary guidance on domestic road access management work, and the manual has been updated to second edition, which has improved some contents in the first edition [3]. At the same time, states introduced regional access management manual combining with local conditions. Lots of broad scholars have studied access management. Williamson et al. [4] studied the access sensitivity geometry index of different road levels to let the geometry design more suitable for vehicle and improve the safety level on the range of access. Schultz [5] also studied the utility of access management in different levels of road and so on. Related research in China started late and has not yet formed a systematic research results and practice.

A lot of researches show that the situation on the urban fringe villages in China is not optimistic. As the core technology in the access management part, access geometric design is lack of normative guidance, leading to the random values of access key geometric parameters and lots kinds of access geometry, which results in much traffic safety hidden danger and not adapting to the development and change of each side of the land and the main line traffic.

This paper is devoted to geometry design method of the level and cross-sectional plane on the urban fringe villages and low-grade road section. This paper chose the urban fringe villages and low-grade road section as the research object to effectively reduce all kinds of traffic conflict and the risk of accidents and enhancing traffic safety level within the range of access as the target, studying access design speed on the urban fringe villages and low-grade road section and proposing the reasonable design values of the key geometric parameters of access level and cross-sectional plane combining domestic current specification and standard. At the same time, access geometric design improvement scheme has been put forward with the demonstration sections and verify the research results and improvement with the method of vehicle trajectory simulation.

Methodology

In the geometric design of the access, the key plane geometry parameters should be given reasonable design values, so that the vehicle can drive in and out of the access safely and smoothly with appropriate speed entering and reducing the conflict severity and accident risk with the range of access. At the same time, pay attention to the security of the pedestrians and non-motor vehicle drivers to improve the traffic safety level.

Sixteen major accesses on the typical urban fringe villages and low-grade road section was chosen to investigate, observing the each side of the spot speed 200 and 15 m disport from the upstream and studying the characteristics of velocity variations of the through vehicles on the main roads. The key geometric parameters influencing the entering and exiting speed can get by measuring the main geometric design parameters, observing the average speed of entering and exiting, analyzing the minimum speed of entering and exiting and the key geometric parameters values.

To obtain the velocity difference in through and entering and exiting access vehicle, reducing serious conflict ratio with the range of access and the casualties risk of pedestrians and non-motor vehicle driver was targeted and the relationship between serious conflict ratio and the velocity difference was searched by a method of observing for a long time. According to related research, by controlling the probability of serious injury of pedestrians and non motor vehicle driver effectively, doing a logistic regression between collision speed and the probability, proposing the suggestion of design speed limit control. Based on the above two points, putting forward the suggestion of design speed and the key geometric design parameters of the level and cross-sectional plane using the suggestion of design speed.

Data Analysis

The Influencing Factors of Minimum Turning Speed on the Access

The survey of speed along a typical urban fringe of low-grade roads (San Jixian Road, Suichang County, Lishui City, Zhejiang Province) indicated that the speed of main roads mainly distributes between 45 and 70 km/h; at the same time, the phenomenon of overspeeding is common, but speed reduces near the access, averagely dropping to 60–80% of the upstream road speed. Furthermore, the speed accessing from main road is further reduced, the entering speed in different access opening and different turning direction showed significant differences in the characteristics, and the two directions also have different characteristics of velocity distribution (Table 1).

Table 1 Average entering and exiting speed distribution along the San Jixian Road

Access number	Access angle (°)	Exit width (m)	Turning radius(m)	Average speed(km/h)
1	13.5	4.0	2.5	8.2
2	57	4.0	7.5	11.3
3	90	5.6	15.0	23.5
4	17.5	9.1	10.0	12.5
5	58.2	4.5	13.0	13.3
6	23	6.6	15.0	13.2
7	90	5.2	8.5	13.7
8	90	10.9	16.0	22.3
9	17	25.5	20.0	20.2
10	90	5.2	2.5	10.8
11	90	6	9.0	12.5
12	16.3	9.7	8.5	14.1
13	75	9.8	6.0	14.4
14	58.3	4.0	4.5	10.1
15	53	11.0	17.5	21.0
16	47.2	10.5	16.0	21.4

Note The average speed is the direction of minimum speed of entering or exiting the access

Table 2 Standard between the average speed of entering or exiting the access and analysis method level dividing of geometry parameter

Level	Speed	Access angle	Exit width	Turning radius
1	V<10 km/h	$J < 45°$	$K < 5$ m	$B < 8$ m
2	10 km/h $\leq$ V< 15 km/h	$45° \leq J< 60°$	5 m $\leq$ K < 10 m	8 m $\leq$ B < 15 m
3	15 km/h $\leq$ V< 20 km/h	$J \geq 60°$	$K \geq 10$ m	$B \geq 15$
4	V $\geq$ 20 km/h	–	–	–

Select three key geometrical design parameters including access angle, outlet width, and turning radius, and the parameters are divided into three levels while the average velocity into four levels (see Table 2). Analysis of variance was performed to detect the main factors of the average speed is the direction of minimum speed of entering or exiting the access. Variance analysis results are shown in Table 3.

Regulate the factor $(Pr > F) < 0.05$ has a high impact on the index and the factor $(Pr > F) < 0.1$ has a significant impact. As given in Table 3, among the access key geometric design indexes, exit width and turning radius have highly significant influence on the speed of vehicles entering and exiting the access, and access angle also has some influence.

Table 3 Variance analysis results

Factor	*F* value	*Pr > F*
Access angle (J)	3.14	0.1065
Exit width (K)	8.12	0.0150
Turning radius (B)	7.42	0.0187

Influence of Access Speed on Traffic Safety

Reducing the Risk of Serious Traffic Conflict Within the Scope of Access

With longtime video scene investigation and field measurement, choose the serious traffic conflict ratio between motors on four directions within the scope of the *T* access and establish the relationship between the ratio and the vehicle speed. Here, the severity of traffic conflict was measured by TTC (time to collision), referring to American Standard, and the TTC less than or equal to 1.0 s is defined as serious traffic conflict.

The Dragon Valley tea village on San Jixian Road, Suichang County, Lishui City, Zhejiang Province, which belongs to urban fringe villages and low-grade road section and has a high access demand and defects of geometric design, was investigated, as shown in Table 4.

Based on the above data, set a correlation analysis between the serious traffic conflict ratio and the speed difference, as shown in Table 5.

The data show significant correlations with the correlation coefficient greater than 0.9. From the above data, the serious traffic conflict ratio increases rapidly with the increase of speed difference. Usually, the serious traffic conflict ratio increases rapidly when the speed difference is greater than 20 km/h.

Reducing the Accident Risk of Pedestrian and Non-Motorized Vehicle in the Area of Access

The death risk of pedestrians and cyclists is closely related to the speed of collision vehicle; and the higher the speed of the vehicle is, the greater the damage to

Table 4 Serious traffic conflict ratio and the speed difference in Dragon Valley tea access

Speed difference (km/h)	Serious traffic conflict ratio (%)			
	Main straight—turn right into the access	Main straight—turn left into the access	Main straight—turn right out the access	Main straight—turn left out the access
≤ 10	3.1	2.8	4	3.8
10–20	7.2	8.3	7.7	6.3
20–30	33.3	35.7	45.5	41.7
30–40	93.8	90	96.4	85.7

Table 5 Derivation analysis on the correlation formula of the serious traffic conflict ratio and the speed difference

Direction	Correlation formula	Derivation correlation formula	Speed difference (km/h)	Growth rate
Main straight—turn right into the access	$y = 1.5464e^{0.1176x}$	$y = 0.1819e^{0.1176x}$	10	0.590
			20	1.911
			30	6.194
			40	20.080
Main straight—turn left into the access	$y = 1.5393e^{0.1187x}$	$y = 0.1827e^{0.1187x}$	10	0.599
			20	1.962
			30	6.431
			40	21.075
Main straight—turn right out the access	$y = 1.9913e^{0.1132x}$	$y = 0.2254e^{0.1132x}$	10	0.699
			20	2.169
			30	6.727
			40	20.866
Main straight—turn left out the access	$y = 1.8017e^{0.1124x}$	$y = 0.2025e^{0.1124x}$	10	0.623
			20	1.917
			30	5.900
			40	18.156

pedestrians and cyclists will be. With the mixed various modes of transportation on the urban fringe villages and low-grade road section, improving traffic safety should be considered emphatically in the protection of pedestrians and non-motor vehicle drivers and other vulnerable groups, so the access design speed must satisfy the damage at lower levels when accidents happening on the pedestrians and non-motor vehicle drivers. This paper draws lessons from the research results of Nie Jin, examining the collision speed to pedestrians and cyclists casualties of traffic accidents, a total of 438 cases [6]. The research results have good applicability because the research data are collected from the whole country, and the sample size is large and the data are weighted according to the national road traffic accident statistics

A regression analysis model of collision speed and pedestrian and cyclist injuries risk is established based on logistic regression analysis. In the formulation, V is the vehicle collision speed; parameters α and β are obtained by maximum likelihood estimation and conducted by statistical tests. The confidence interval is 95%, and the regression equation of vehicle collision speed and pedestrian injury risk (probability) is as follows:

$$P(v) = \frac{1}{1 + \exp(2.991 - 0.083v)} \tag{1}$$

In the same way, the regression equation of the vehicle collision speed and the serious injury risk (probability) is as follows:

Table 6 Results of logistic regression analysis

	α	β	The value of P
Pedestrian sample	2.991	−0.083	<0.0001
Rider sample	4.343	−0.074	<0.0001

The Risk of Pedestrian Injury
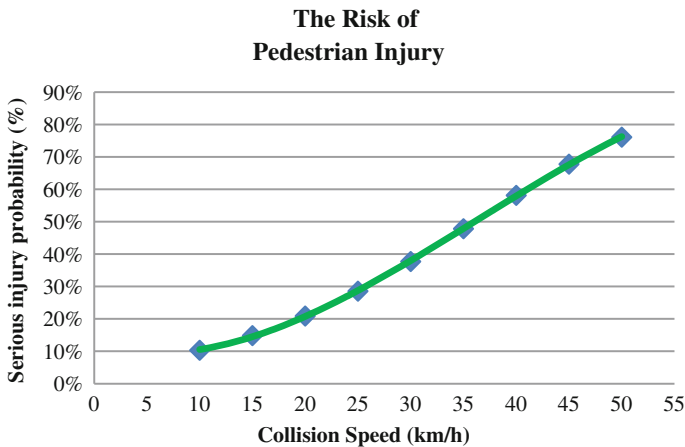
Fig. 1 Relationship between the risk of pedestrian serious injury accident and the speed of collision

$$P(v) = \frac{1}{1 + \exp(4.343 - 0.074v)} \tag{2}$$

(1) and (2): V is vehicle collision speed and the unit is km/h.

All of the above two are tested by chi-square test, and P is less than 0.0001, which indicates that the collision speed of vehicle is significantly related to the risk of serious injury of bicycle and rider. The results of logistic regression analysis in this study are shown in the following table (Table 6).

From Figs. 1 and 2, pedestrians have the higher probability of crash injuries compared to non-motor vehicle drivers, and the injured probability of pedestrian or non-motor vehicle driver is improved rapidly after the collision speed is greater than 30 km/h.

Study on Access Design Speed

According to the analysis of the data, the following conclusions can be drawn.

- In fringe villages and low-grade road section, the access design speed should not be more than 30 km/h; otherwise, the travel environment of pedestrians and non-motor vehicle drivers will be in a great risk.
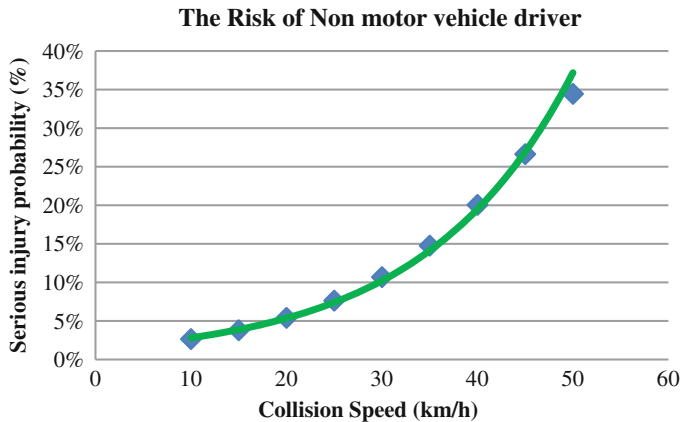
Fig. 2 Relationship between the risk of non-motor vehicle driver serious injury accident and the speed of collision

- To reduce serious traffic conflict ratio and accident rate, the vehicle speed difference between through vehicle on the main road and turning vehicle in every direction within the access should be within 20 km/h; if the condition is limited, the speed difference should not exceed 30 km/h.

According to the above conclusions, considering China's current highway design standards, access design speed in the fringe villages and low-grade road section is proposed as follows:

- The car turning design speed value should be taken for 15–25 km/h; when condition is better, design speed value should not be higher than 30 km/h; if the condition is limited, design speed should not be less than 10 km/h.
- The turning design speed value of large passenger cars, trucks, articulated buses, and articulated train should be taken for 15–20 km/h; when condition is better, design speed should not be higher than 25 km/h; if the condition is limited, design speed should not be less than 10 km/h.

Case Study of Improving Access Geometry Design

Situation of the Access and Security Risk Analysis

Based on the former research achievements, key access on the San Jixian Road in Suichang County, Lishui City, Zhejiang Province, was chosen to do geometry design safety improvement. The access is Dragon Valley tea Xiang Pai Fang access, located at the junction of Zhen Da Zhe and Yongan, the access road skews with the

Fig. 3 Location of Dragon Valley tea arch access

main road (San Jixian Road) in a small angle, and the specific situation is depicted in Fig. 3.

The key geometric parameters of the access and traffic operation are as follows, shown in Tables 7 and 8.

According to the geometrical dimensions of the access port and the operation of the traffic, the hidden safety danger of the access is analyzed:

- The access angle is only 23°, the small angle leads to the turning radius too low on part of the directions. The speed difference between the turning speed and the straight driving speed is too large to form a large security risk. In another direction, the low angle of the intersection with the main road leads to the speed of some turning right vehicle on the access road too large, making the pedestrians and non-motorized vehicles in a great threat;
- With the obstacle of sight, such as the vegetation and arch, the risk of accident increases.

Design of Access Geometry Improvement Scheme

According to the problems existing in the access, the following geometric design improvement scheme is proposed, see Fig. 4.

The improvement scheme eliminates the small angle crossing, and the access angle is promoted from 23 to 60°, which is in accordance with the relevant provisions in the current code [7, 8]. The access road plane alignment on the access is partly improved, and the speed difference between through vehicles on the main road and vehicles in and out the access is control by setting the circular curve

Table 7 Key geometric parameters of Dragon Valley tea arch access

Access angle	23°	Access path slope	5%(Downhill access to the main road)
Access path length	>200 m	Access path width	5.5 m
Service land	Residential, commercial, administrative office	Obstacle of sight	Botany
Small angle turning radius	<10 m	Speed reduction facility	With an export slowdown ridge

Table 8 Traffic operation of Dragon Valley tea arch access

Main road traffic volume	416 veh/h
Access road traffic volume (in, out)	24 veh/h
Main road (access upstream 200 m) average speed (East–West)	50.8 km/h
Main road (access upstream 200 m) average speed (West–East)	51.2 km/h
Main road (access upstream 15 m) average speed (East–West)	35.4 km/h
Main road (access upstream 15 m) average speed (West–East)	34.6 km/h
The average speed of vehicles on the main road turning right into the access	12.1 km/h
The average speed of vehicles on access road turning left out of the access	14.3 km/h
The average speed of vehicles on small radius turning direction	13.2 km/h
The average speed of vehicles on the main road turning left into the access	30.7 km/h
The average speed of vehicles on access road turning right out of the access	34.3 km/h
The average speed of vehicles on big radius turning direction	32.5 km/h
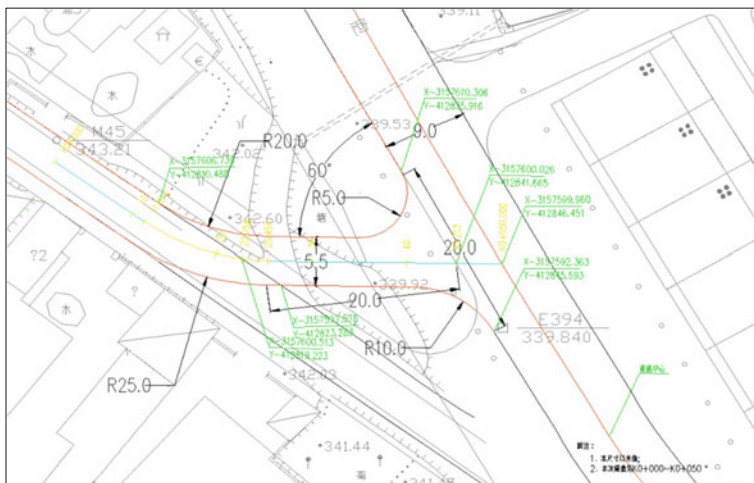
Fig. 4 Improvement of Dragon Valley tea access

(lateral radius of 25 m, inner radius of 20 m) to control high exit speed in the design speed within the access described in Sect. 3.3. The length of the straight line behind the circle curve is designed to be 20 m, and the slope is not more than 8%, which can meet the braking requirements of an articulated vehicle. In addition, curb radius at the outlet can be improved correspondingly, which can guarantee the vehicle in and out access with the design speed.

Evaluation of Geometric Design Improvement Scheme for Access

By using the vehicle turning design software AutoTURN for traffic simulation, which is developed based on AutoCAD platform, the various types of vehicles

Fig. 5 **a** Vehicles on the main road turning *right* into the access, **b** vehicles on the main road turning *left* into the access, **c** vehicles on access road turning *right* out of the access, and **d** vehicles on access road turning *left* out of the access

vehicles on the main road turning right into the access

vehicles on the main road turning left into the access

(c)

vehicles on access road turning right out of the access

(d)
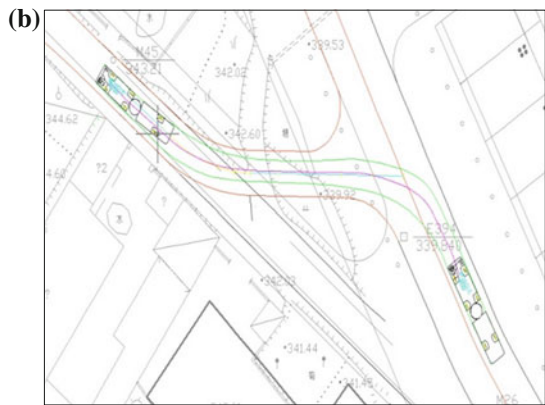
vehicles on access road turning left out of the access
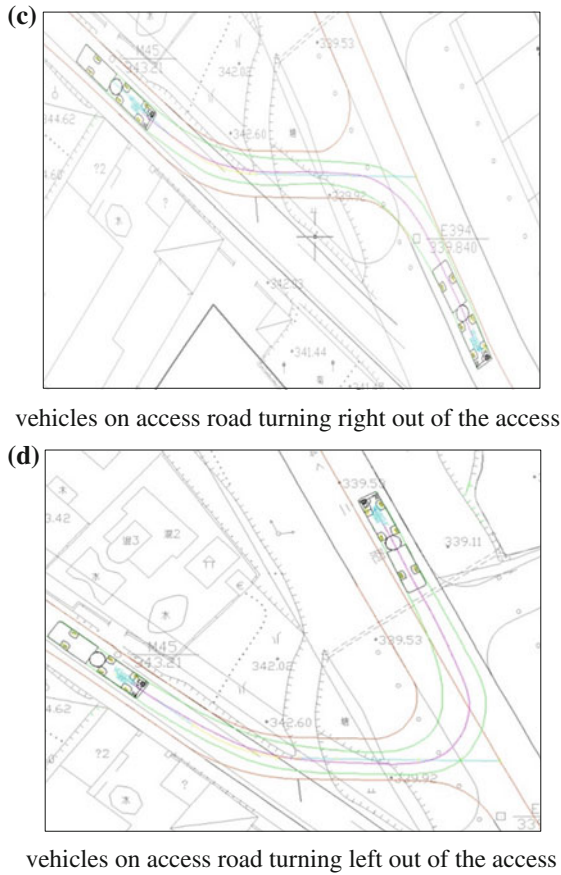
Fig. 5 (continued)

driving track can be simulated and have a good application on the road traffic design.

Because of the complex land use components he famous Dragon Valley access services having residential, commercial, administrative office and other functions, the articulated buses is chosen as design vehicle on the access. Taking into account the needs of traffic safety, the turning speed on the access and the speed of through vehicle are controlled in 20 km/h, and the vehicle turning speed is set to 15 km/h in the simulation.

According to the four directions such as vehicles on the main road turning right into the access, vehicles on the main road turning left into the access, vehicles on access road turning right out of the access, and vehicles on access road turning left out of the access, the design vehicle driving tracks were simulated, corresponding to the trajectory as shown in Fig. 5a–d.

In the four pictures above, the red dashed line is the edge of the road line, the green line is vehicle trajectory line, and the powder solid line is vehicle center line. From the pictures, turning trajectory design of the vehicles is in the range of the width of the road, and the wheels are not been separated from the cement concrete pavement into soil shoulder. In addition, considering the stalls set up or stacked items on both sides of fringe villages and low-grade road section, road design should set aside a certain safety redundancy outside the vehicle trajectory as far as possible. In the pictures, except for design vehicle on the main road turning right into the access with the preset speed, ensure design vehicle in the remaining three directions driving having no less than 0.5 m distance space between the wheels and the road edge. Even on the situation of vehicle on the main road turning right into the access, vehicle track is access parking line position near the pavement edge (where the vehicle trajectory is not beyond the reach of the pavement) while in other areas of the access, the direction of the vehicle driving can still maintain safe distance with the pavement edge enough.

Conclusions and Prospects

For typical traffic environment on the fringe villages and low-grade road section, key geometric parameters influencing the speed of vehicles in and out the access were analyzed; combining with the collection and analysis of traffic conflict data and to the research of pedestrians and non motor vehicle driving traffic accidents data, the access design speed on the fringe villages and low-grade road section is put forward. According to the typical access for geometric design improvement, using the vehicle trajectory simulation method, that the access design speed suggestion is reasonable and the geometric improvement design is effective was indicated. The future research should also strengthen the aspects of the longitudinal geometric design, and sight protection in order to form a complete guidance system for the geometric access design on the fringe villages and low-grade road sections.

References

1. Conggao, L., and W. Yingna. 2007. Road traffic safety analysis and countermeasures of China's urban and rural fringe. *Traffic Standardization* 12: 180–182.
2. Zhonghua, W., C. Jia, and C. Yongsheng. 2010. The urban fringe of Beijing city department of highway of high risk areas of origin analysis. *Journal of Beijing University of Technology* 36 (2): 223–227.
3. Committee on Access Management. 2003. *Access management manual*. Washington D C: Transportation Research Board.
4. Michael Williamson, Mohammad Jalayer, Huaguo Zhou, Mahdi and Pour Rouholamin. 2014. A sensitivity analysis of crash modification factors of access management techniques in

highway safety manual. In *2nd International Conference on Access Management: Access Management Theories and Practices*, U.S.

5. Schultz Grant G, Braley Kordel T, and Boschert Tim. 2009. Prioritizing access management implementation. *Transportation Research Record* 57–65.

6. Nie Jin, W., J.W. Lingtao, R. Lihai, and Y. Jikuang. 2014. The city road speed of pedestrian and bicycle traffic safety value of China. *Journal of Highway and Transport, Based on the Transport* 27 (7): 91–97.

7. The People's Republic of China industry standard. 2006. *D20-2006 JTG Highway Route Design Code*. Beijing: People's Communications Press.

8. The People's Republic of China industry standard. 2014. *D20-2014 JTG Highway Route Design Code*. Beijing: People's Communications Press.

Stochastic Traffic Assignment Model Considering Park & Ride Network and Travel Time Reliability

Anning Ni, Xuxun Lin and Jing Luo

Abstract Most existing state-of-the-art combined-modal traffic assignment models seldomly consider the influence of travel time reliability on travel behaviors. In order to reveal the influence of travel time reliability on travelers' route and mode choice behaviors, this paper constructs a combined inter-modal assignment model based on travel time reliability, which takes travel time and travel time reliability jointly constitute travel cost. The solution is then designed based on K-shortest paths algorithm coupled with successive average method. The validation of the proposed model and algorithm of solution are conducted using a real road network of Shanghai North Bund. The results are compared with that of the traditional models without consideration of the impact of travel time reliability. In addition, this research analyzes the effect of P&R policy on reducing traffic congestion as well as improving travel time reliability and demonstrates the three-dimensional relationship among the proportion of travel demand using two alternative modes, reliability preference factor and travel demand utility factor.

Keywords Stochastic traffic assignment · Travel time reliability · Combined-modal · Elastic demand

A. Ni · J. Luo
School of Naval Architecture, Ocean and Civil Engineering, Shanghai Jiao
Tong University, 800 Dongchuan Road, Shanghai 200240, China
e-mail: nianning@sjtu.edu.cn

J. Luo
e-mail: carryonex@sjtu.edu.cn

X. Lin (✉)
Business School Changzhou University, 1 Middle Gehu Road,
Changzhou 213164 Jiangsu, China
e-mail: linxuxun2002@gmail.com

Introduction

Conventional traffic assignment models generally adopt average travel time as the determinant of route choice regardless of the fluctuation of network supply and travel demand level. However, it could not be neglected that a traveler's decision of route choice is invariably affected by various random factors under actual conditions, which implies that an individual's experienced travel time is a random variable. Furthermore, studies have shown that travelers take not only the length of travel time into account, but travel time reliability as well when they make route choice decisions. There is a need to consider travel time reliability in the models of travelers' behavior of route selection. Consequently, researchers have proposed a number of traffic assignment models integrating travel time reliability issue. The supply and demand of traffic network are stochastic in nature. Moreover, the unexpected events and incidents in traffic network can bring other kinds of variability. Drivers are not able to anticipate these variabilities, which may result in additional errors in their route selection. The variability of highway travel times has become a concern for travelers as well as transportation management agencies. Travel time reliability, as the opposite of variability, has been treated as an important key factor for travelers to secure their on time travels. At present, the developing of reliability and variability measures has been given much attention and effort. The current measures of travel time reliability include the standard deviation of travel time, the coefficient of variation of travel time, the planning-time index, the travel time index, the buffer index, the 90th or 95th percentile travel time, and the misery index. Using these travel time reliability measures, researchers have conducted a lot of studies in traffic assignment modeling with the integration of travel time reliability.

Literature Review

Mirchandani used the variance of random travel time as reliability measure (*1*). He proposed a stochastic user equilibrium traffic assignment model based on stochastic travel time, supposing that the costs of travelers were comprised of both travel time and travel time reliability. On this base, Liu constructed a dynamic traffic assignment model using variational inequality (*2*). Assuming all travelers tend to be risk-averse, Srinivasa introduced the fluctuation of travel time into link travel cost function and constructed an user equilibrium (UE) traffic assignment model (*3*). A comparison with traditional UE model is made as well. Using field data to fit route travel time stochastic distribution function, Joel P. Franklin associated early \late arrival penalty factors with travelers' utility function to provide a model depended on time-based window and optimal departure time selection (*4*). In addition, a case study was carried out in Stockholm city of Sweden. Anthony Chen and Zhongzhou proposed A-beyond Average Value reliability traffic assignment model (*5*). They took the conditional expectation of travel time exceeding the estimated value as travelers' minimum objective, which simulated travelers' route

choice behavior under both reliability and non-reliability conditions more deeply. Zhongzhou et al. compared results of A-beyond Average Value reliability traffic assignment model with that of other two conventional traffic assignment models under the condition of stochastic traffic demand, which obtained some results of influence of different cognition toward risks and preference on route choice (*6*). Shao built a traffic assignment model on the basis of stochastic travel demand and proposed the cost calculation index of travel time budget (*7*). In Shao's model, travel time was represented by linear combinations of mean-variance, which can simulate travelers' route choice behavior under the reliability condition. ManWong et al. put the model forward by introducing a traffic simulation model based on cellular automaton and revealed dynamic change process of travel time reliability (*8*). Siu Wing Yee presented a traffic assignment model on the basis of travel time budgeting theory and assessed some traffic management measurements regarding the random fluctuation network such as congestion charge and increasing traffic capacity. A dynamic model was then constructed to reveal the influence of random fluctuation of road network on travelers' departure time decisions (*9*). However, due to the high complexity of the analytic model and algorithm, only one origin–destination pair and two routes were considered. Allowing for the uncertainties of travelers' departure time decisions caused by the random fluctuation of travel time, Hao Li et al. established dynamic simulation model to study the correlativity between various road network structures and random distribution function of departure time (*10*). Szeto et al. developed a route choice model for multiple users based on reliability with considering individuals' diverse preference toward risk (*11*). They also conducted research on how the travelers' perceived risk costs affect the travel time reliability of road network through experimental study.

The above existing studies have proposed a number of reliability measures from various points of view to comprehensively analyze the influence of random fluctuation of network supply and demand on travel behaviors. However, the following several factors have not been investigated sufficiently yet in most existing models: (i) Most models are inelastic demand models which do not get the demand change associated with the network travel cost involved. (ii) Merely a sole travel mode is considered whereas the mutual influences among different modes are neglected much less in terms of travelers' mode and route choice decisions caused by network random fluctuation under the combined-mode condition. However, with the gradually completeness of traffic infrastructures, combined-mode has been one of the common optional modes to metropolis travelers. (iii) The known traffic assignment models just provide mathematical example analyses, which are not put into practical road network applications. For these reasons, there exists potential improvement in current studies. This paper proposes reliability-based combined-mode random elastic demand traffic assignment model. This study focuses on how does combined-mode affect travelers' mode choice and route choice behavior in the context of the network that the capacity of links fluctuates randomly. Furthermore, on the basis of real networks, this paper analyzes the function of P&R policy performing on the improvement of network reliability and the change in reliability preference factor influencing on demand proportion of different modes.

Traffic Assignment Model Based on P&R and Network Reliability

This paper constructs a traffic assignment model with elastic demand based on travel time reliability including two travel modes: (i) travel by motor vehicle; (ii) travel by P&R for subway. Considering a traffic network $G = (N, L)$, N and L denote the network nodes set and links set connecting between two nodes. R and S are starting (terminal) points set where $R \subset N$, $S \subset N$, respectively. Furthermore, it denotes P&R stops set $I \subset N$ and M denotes the choice set of travel modes, $M = \{a, b\}$ a and b are, respectively, corresponding to the two kinds of modes in aforesaid setting condition. W is the selectable path set of each OD pair that defines by w_{rs}^m which means the mode m is chosen between the OD pair of r and s. The route travel time equation can be written as follows:

$$h_{p_k} = \sum \delta_{a,p_k} t_a(x_a) \tag{1}$$

where

h_{p_k} = the travel time of link p_k
$x_a = \sum \delta_{a,p_k} f_k$, f_k is the flow of the link p_k
δ_{a,p_k} = a dummy variable
k = a serial number.

Thus

$\delta_{a,p_k} = 1$, if the link a is a part of the route p_k;
$\delta_{a,p_k} = 0$, otherwise

Travel Time Reliability Evaluation of Motor Vehicle Link

Assuming the traffic capacity c_a is a uniformly distributed random variable, $\overline{S_a}$ indicates the design traffic capacity corresponding to $\underline{S_a}$ of the capacity when congestion. Thus if taking $\underline{S_a} = \beta^a \overline{S_a}$ and $\beta^a \in (0, 1)$, it can be obtained that c_a obeys uniform distribution on $[\beta^a \overline{S_a}, \overline{S_a}]$. In this study, the link travel time reliability is defined as a probability that the real link travel time less than a certain given threshold.

$$R_a = P(t_a < T_a) \tag{2}$$

where

R_a = travel time reliability of link a;
t_a = real travel time of link a;

T_a = travel time threshold; and $T_a = kt_a^0 (k > 1)$, which generally takes the multiple of travel time when zero flow rate.

Moreover, BPR function is usually used to calculate the link travel time and can be expressed as:

$$t_a = t_a^{(0)}[1 + 0.15(d_a/c_a)^4] \tag{3}$$

Travel time reliability of the link a can be formulated by a continuous subsection function according to the uniform distribution expression when the BPR function is substituted to Eq. (3):

$$R_a = \begin{cases} 1 & m_a \leq \underline{S_a} \\ \int_{m_a}^{\overline{S_a}} f_{C_a}(t)\mathrm{d}t & \underline{S_a} < m_a < \overline{S_a} \\ 0 & m_a \geq \overline{S_a} \end{cases} \tag{4}$$

where

$f_{C_a}(t)$ = probability density function of uniform distribution
$m_a = \frac{x_a}{[(k-1)/\alpha]^{1/\beta a}}$

Calculation of Travel Time Reliability Refers to the Transfer Between Motor Vehicle and Subway

In this study, it is presumed that the transfer time between motor vehicle and subway consists of the time spending on looking for a parking place, denoted as T^s, and driver's waiting time in the subway station, denoted as T^w.

Time spending on looking for a parking place

T_j^s is taken as average time of traveler looking for a parking at P&R place j. According to the literature (*12*), the time of finding a parking place can be formulated as:

$$T_j^s = d_j(v_j) = d_j^0 + 0.31(v_j/C_j)^{4.03}, \quad \forall \, j \in I_{rs} \tag{5}$$

where

d_i^0 = the time cost of finding parking place when there is no vehicle in the garage;
v_i = the demand of car parking;
C_i = the total number of parking place in parking lot.
Driver's waiting time in the subway station
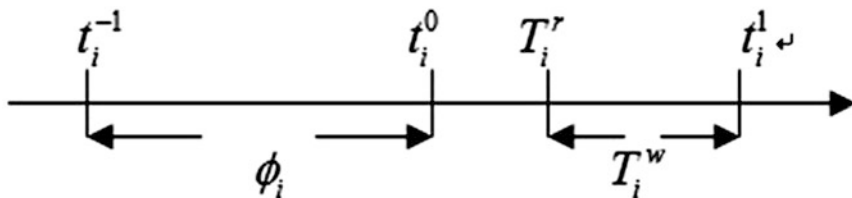
Fig. 1 Illustration of subway waiting time

Let ϕ_i be the departure interval of the subway line 'i'; t_i^0 be the leaving time of the traveler's target subway; t_i^{-1} be the leaving time of previous subway; and t_i^0 be the next one. Due to the time interval of subway traveling between two fixed stations is relatively stable, it supposes that the subway arriving time to destination is a constant denoted as T_i; the traveler arriving time to the subway station is denoted as T_i^r; and the traveler waiting time at the subway station is T_i^w. According to the research of the literature (13), travelers whose choice of subway line I arriving at the station of time T_i^r obey some kind of distribution of mean value t_0 on the interval of $[t_i^{-1} t_i^1]$, which is shown in Fig. 1.

$$T_i^w = \begin{cases} t_i^1 - T_i^r & T_i^r > t_i^0 \\ t_i^0 - T_i^r & T_i^r < t_i^0 \end{cases} \tag{6}$$

Travel time reliability of subway line can be written as following formula:

$$R_i^b = P(T_i + T_i^w + T_j^s < kT_i | T_i^r < t_i^0) * P(T_i^r < t_i^0) + P(T_i + T_i^w + T_j^s < kT_i | T_i^r > t_i^0) \\ * P(T_i^r > t_i^0) \tag{7}$$

Substituting Eq. (7) into Eq. (6) and using the conditional probability formula:

$$R_i^b = P(t_i^0 + T_j^s + (1 - k)T_i < T_i^r < t_i^0) + P(T_i^r > \max(t_i^0, t_i^1 + (1 - k)T_i) + T_j^s) \tag{8}$$

Function of Travel Cost

Travel cost of link a is denoted by C_a including the notion of both travel time and travel time reliability:

$$C_a = t_a/(1 + \lambda R_a) \tag{9}$$

λ represents the reliability preference factor, which reflects travelers' relative preference level.

Modeling Structure

In reality, travelers just can roughly estimate the link travel cost, so that different travelers would give different estimations. The differential between perceptual value and actual value of travelers' route travel cost is usually assumed as a random variable. When the network flow assignment is satisfied with stochastic user equilibrium conditions, travelers could not reduce their perceptual cost by changing routes.

Let $V_{rs,p}^m$ be perceptual travel cost of OD pair between r and s by travel mode of m and route p, then it can be obtained:

$$V_{rs,p}^m = C_{rs,p}^m + (1/\theta_{rs,p}^m)\xi_{rs,p}^m \tag{10}$$

In this equation, $\xi_{rs,p}^m$ corresponding to random error term and $\theta_{rs,p}^m$ is a constant which acts as a role of transferring travel cost to the utility.

It supposes that the probability of choosing route p is $P_{rs,p}^m$ when traveler's choice of mode m is known, then:

$$P_{rs,p}^m = P(C_{rs,p}^m < C_{rs,i}^m, \quad \forall\ i \in w_{rs}^m) \tag{11}$$

In the case of combined-model, traveler should do the mode choice decision before the route choice. Generally speaking, the probability of traveler's mode choice satisfies the logit model expression:

$$q_{rs}^m = Q_{rs} \frac{\exp(-\mu_{rs}^m h_{rs}^m)}{\sum_m \exp(-\mu_{rs}^m h_{rs}^m)}, \quad \forall\ r \in R, s \in S, m \in M \tag{12}$$

h_{rs}^m denotes the minimum travel cost of mode m, i.e.,:

$$h_{rs}^m = -(1/\theta_{rs,p}^m) \ln \left(\sum_{p \in P_{rs}^m} \exp\left(-\theta_{rs,p}^m C_{rs,p}^m\right) \right) \tag{13}$$

μ_{rs}^m is a constant, which corresponds to the travel cost-utility conversion coefficient and satisfies the condition of $0 < \mu_{rs}^m \leq \theta_{rs,p}^m$.

The minimum expected travel cost between the OD pair r and s can be expressed as following pattern:

$$C_{rs} = (-1/\mu_{rs}^m) \ln \left(\sum_m \exp(-\mu_{rs}^m h_{rs}^m) \right) \tag{14}$$

Lastly, this paper takes the elastic demand between OD pair into account. $D_{rs}()$ indicates the traffic demand function of OD pair between r and s. Then $D_{rs}^{-1}()$ is the demand inverse function, thus

$$Q = D_{rs}^{-1}(C_{rs}) \tag{15}$$

The demand inverse function in this paper adopts following exponential form:

$$Q_{rs} = q_{rs}^0 \exp(-\rho C_{rs}) \tag{16}$$

q_{rs}^0 represents the maximum potential traffic volume between the OD pair of r and s. ρ is a constant, which reflects the sensitivity of the traffic volume to the expected travel impedance. According to above assumed conditions, the combined-mode traffic assignment model based on travel time reliability can be expressed as following variational inequality form:

$$\sum_{rs} \sum_{m \in M} \sum_{p \in p_{rs}^m} C_{rs,p}^m (f_{rs,p}^m - f_{rs,p}^{m*}) + \sum_{rs} \sum_{m \in M} \frac{1}{\mu_{rs}^m} \log(q_{rs}^{m*})(q_{rs}^m - q_{rs}^{m*})$$

$$+ \sum_{rs} \sum_{m \in M} \sum_{p \in p_{rs}} \frac{1}{\theta_{rs,p}^m} \log\left(\frac{f_{rs,p}^{m*}}{q_{rs}^{m*}}\right)\left(f_{rs,p}^m - f_{rs,p}^{m*}\right) - \sum_{rs} D_{rs}^{-1}(Q_{rs}^*)(Q_{rs} - Q_{rs}^*) \geq 0 \tag{17}$$

And constraint conditions:

$$\sum_m q_{rs}^m = Q_{rs}$$

$$\sum_{p \in P_{rs}^m} f_{rs,p}^m = q_{rs}^m$$

$$f_{rs,p}^m \geq 0$$

Solution Algorithm

On the basis of K-shortest-path algorithm and MSA iteration convergence algorithm, we provide the solution method by following steps:

Initialization

Setting the initial traffic flow rate of each link as zero, then the minimum expected travel cost of each mode and each OD pair can be calculated by Eqs. (13) and (14). And traffic demand Q_{rs} between each OD pair can be obtained according to Eq. (16) as well. Then the volume of two travel modes is determined by Eq. (10). In the end,

assigning the flow into the random network in terms of each mode, it can obtain a set of feasible links volume, then let $m = 1$ (m represents the iterative times).

Updating the Link Travel Cost

Link travel time and its reliability can be figured out depending on current feasible link traffic flow set. Then travel cost of each link can be calculated by Eq. (9).

Determining the Search Direction

On the basis of present link travel cost, conducting flow assignment of random network, it can obtain the additional traffic volume $y_a^{(m)}$ of link a and the search direction $d_a^{(m)} = y_a^{(m)} - x_a^{(m)}$.

Updating Link Traffic Flow

Utilizing the step of $1/m$ set in advance, the link traffic flow can be conducted for update computation: $x_a^{(m+1)} = x_a^{(m)} + (1/m)(y_a^{(m)} - x_a^{(m)})$

Convergence Criterion

Convergence criterion '$\min_a |x_a^m - x_a^{m-1}| < \varepsilon$' is employed with the error term ε. If the previous setting error condition is satisfied, stop; otherwise, let $m = m + 1$ and back to step (1).

In the above algorithms, random network flow assignment requires to be executed in every iteration process. The random network flow assignment procedure in the algorithm would be displayed as follows:

1. Computing the travel cost of previous K-shortest paths between all the OD pairs through the present link travel cost.
2. Computing the choice probability of each route

$$P_i = e^{C_i} \Big/ \sum_{j=1}^{K} e^{C_j}, \quad i = 1, 2, \ldots, j, \ldots K$$

In the end, additional traffic volume $y_a^{(m)}$ of each link can be calculated by the route-link matrix.

Empirical Application

For the sake of verifying model and algorithm efficiency, it takes a certain network of Shanghai North Bund as an example to conduct the analysis. As shown in the Fig. 2, this network is totally comprised of 19 motor roads and a subway line (dotted line).

The travel time on free flow condition and the traffic capacity of each link are shown in the Tables 1 and 2. And the pcu in the table denotes passenger car unit.$k = 1.1$, $\theta_{rs,p}^m = 0.20$, $\mu_{rs}^m = 0.10$, $\lambda = 0.5$, $\rho = 0.01$, in addition $\beta = 0.2$ for all links. Considering nodes 1–3, 1–4, 2–3, and 2–4 of the OD pair, they correspond to the potential traffic volume of 600, 500, 500, and 600 passengers/h, respectively.

Dotted line represents the 12th subway line under planning, which would cost 7 min in this area with departure interval ϕ of 2 min. According to Eq. (5), it builds a P&R transfer point at the place where the subway enters this network zone and supposes T_i^r obeys the uniform distribution on $[t_i^{-1}t_i^1]$ with the mean of t_0, and $d_j^0 = 0.5$, $C_j^0 = 2000$. Then through Eqs. (6)–(8), the reliability of subway can be obtained as one.

Utilizing the model and algorithm proposed by this paper, total demands of two modes at the equilibrium condition are 1606 p/h and 794 p/h, respectively. Link performance indices are shown in Table 3.

Fig. 2 Shanghai North Bund network structure

Table 1 Link ID–Name relation

ID	Name of link	ID	Name of link
1	Haining Rd.	11	Kunming Rd.
2	Zhoujiazui Rd. (West)	12	Dalian Rd. (Middle)
3	Zhoujiazui Rd. (East)	13	Wusong Rd. (South)
4	Wusong Rd. (North)	14	Xinjian Rd. (South)
5	Xinjian Rd. (North)	15	Gongping Rd. (South)
6	Gongping Rd. (North)	16	Dalian Rd. (South)
7	Dalian Rd. (North)	17	Changzhi Rd. (West)
8	Hanyang Rd.	18	Changzhi Rd. (East)
9	East Hanyang Rd.	19	East Changzhi Rd.
10	Tangshan Rd.		

Table 2 Link data

Link	$t_a^{(0)}$ (min)	c_a (pcu/h)	Link	$t_a^{(0)}$ (min)	c_a (pcu/h)
1	1.33	3588	11	1.85	500
2	0.88	3563	12	0.20	3626
3	1.28	3563	13	0.33	4743
4	0.66	4743	14	0.39	242
5	0.63	242	15	0.30	773
6	1.02	773	16	0.24	3626
7	0.35	3626	17	2.27	1654
8	1.94	892	18	1.23	1654
9	1.31	892	19	1.87	929
10	1.86	892			

Comparing the link performance indices to those without P&R policies, results are given in Table 4.

R1 and R2 refer to the reliability indices before and after the P&R policies implemented. Although total trips still increase, it can be seen that reliability of the majority of the links is improved after P&R polices taken effect, except link 9 and 16. Moreover, the number of links that reliability equal to zero decreased dramatically due to the subways appeal to a part of travelers, which result in the road traffic volume reduced and the travel time reliability improved.

Finally, Fig. 3 shows the three-dimensional relationship among the varying proportion of two kinds of travel modes, reliability preference factor, and travel demand utility factor.

It can be summarize from Fig. 3 that no matter what size of the traveler's utility conversion factor, the quantity demand of park & ride increases along with the increase of reliability preference factor coefficient. Compared to the reliability of motor vehicles, subway would be better. And the increase of reliability preference factor implies that the travelers' reliability preference could increase as well.

Table 3 Link flow, travel time, and travel time reliability

Link	Flow (pcu/h)	Reliability	Link	Flow (pcu/h)	Reliability
1	760	0.9571	11	405	0.1293
2	376	1	12	708	0.9797
3	154	1	13	159	1
4	555	1	14	160	0.3364
5	242	0.01	15	173	0.9401
6	274	0.7602	16	686	0.9881
7	154	1	17	159	1
8	777	0.0451	18	342	0.9643
9	860	0.01	19	492	0.518
10	638	0.2612			

Table 4 Comparison of link reliability before and after P&R

Link	R1	R2	Link	R1	R2
1	0.8626	0.9571	11	0	0.1293
2	1	1	12	0.8605	0.9797
3	1	1	13	1	1
4	1	1	14	0.0279	0.3364
5	0	0	15	0.8518	0.9401
6	0.5928	0.7602	16	1	0.9881
7	1	1	17	1	1
8	0	0.0451	18	0.8381	0.9643
9	0	0	19	0.2227	0.518
10	0	0.2612			

Fig. 3 Three-dimensional map of travel amount ratio according to the change of λ and μ
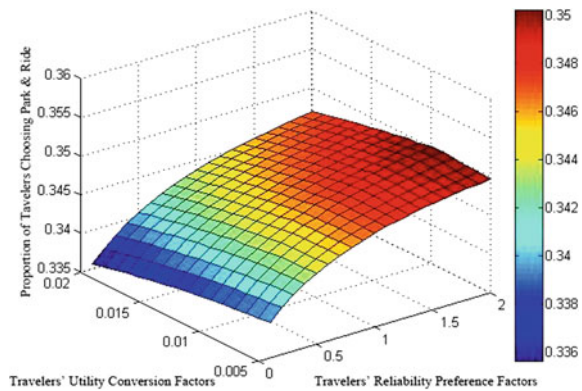
Consequently, no matter what size of the travelers' utility conversion factors, the proportion of travelers choosing park & ride would go rise, which accords with the actual situation.

Conclusion

Current combined-mode models mostly fail to get the influence of reliability factor involved. To solve this problem, reliability-based combined-mode traffic assignment model is proposed to better reveal the mode and route choice behaviors, which allows for the network random fluctuation on the condition of combined-mode travel. This paper designs the mode and flow assignment algorithm based on the K-shortest paths and MSA iteration convergence algorithm. Model efficiency and accuracy are verified by a section of practical network. The research also pays a significant attention on how combined-mode travel influences on travel mode and route choice behavior in a network where the traffic capacity fluctuates randomly. In addition, it analyzes the role of P&R policies on improving the network reliability. Finally, it provides the three-dimensional relationship among the proportion of travel demand of two modes, the reliability preference factor, and the travel demand utility factor. However, the issue of traffic demand fluctuation effect is not covered in this paper, which could be recommended for future studies.

References

1. Mirchandani, P., and H. Soroush. 1987. Generalized traffic equilibrium with probabilistic travel times and perceptions. *Transportation Science* 21: 133–152.
2. Liu, H., X. Ban, and P. Mirchandani. 2002. Analytical dynamic traffic assignment model with probabilistic travel times and perceptions. *Journal of the Transportation Research Board* 1783 (National Research Council, Washington, D.C.).
3. Srinivasa. 2007. *Traffic assignment model with travel time reliability consideration*, vol. 7. El Pasto: University of Texas.
4. Franklin, Joel P. 2009. Travel time reliability for Stockholm roadways: Modeling the mean lateness factor. *Transportation Research Annual Meeting*.
5. Chen, Anthony, and Zhong Zhou. 2010. α-reliable mean-excess traffic equilibrium model with stochastic travel times. *Transportation Research Part B* 44: 493–513.
6. Zhou, Zhong, and Anthony Chen. 2008. Comparative analysis of three user equilibrium models under stochastic demand. *Journal of Advanced Transportation* 42 (3): 239–263.
7. Shao, H., W.H.K. Lam, Q. Meng, and M.L. Tam. 2006. Demand-driven traffic assignment problem based on travel time reliability. *Transportation Research Record: Journal of the Transportation Research Board*, vol. 1985, 220–230. Washington, D.C.: TRB, National Research Council.
8. Ng, Man Wo. 2009. A dynamic route choice model in face of uncertain capacities. In *Transportation Research Board Annual Meeting*, Washington, D.C.
9. Siu, Wing Yee. 2009. *Reliability-based transportation network studies*. Hong Kong University of Science and Technology.
10. Li, Hao. 2009. Modeling departure time choice under stochastic networks involved in network design. In *Transportation Research Board Annual Meeting*, Washington, D.C.

11. Szeto, W.Y, and A. Sumalee. 2009. Multi-class reliability based stochastic dynamic user equilibrium assignment problem with random traffic states. In *Transportation Research Board Annual Meeting*, Washington, D.C.
12. Li, Zhi-chun, W.H.K. Lam, S.C. Wong, Dao-li Zhu, and Hai-jun Huang. 2007. Modeling Park & Ride services in a multi-modal transportation network with elastic demand. *Transportation Research Board Annual Meeting TRB*.
13. Niels van Oort. 2009. Controlling operations of public transport to improve reliability: Theory and practice. In *Transportation Research Board Annual Meeting*, Washington, D.C.
14. Zhou, Jing, and Yan Xu. 2001. Stochastic user equilibrium assignment with elastic demand and its application. *Journal of System Engineering* 16 (2): 88–94.
15. Yuan, Hong-tao, and Mei-zheng Zhu. 2004. A fast algorithm and it's implementation of finding the k-shortest paths. *Computer Engineering and Applications* 40 (6): 51–53.

Inverse Kinematics Model's Parameter Simulation for Stewart Platform Design of Driving Simulator

Gan Hu, Xiaomeng Li and Xuedong Yan

Abstract Due to the wide use of driving simulator in driving behavior and safety investigation as well as the application in drivers training, research about the experience authenticity of driving simulator gains more attention than ever. This paper focuses on the parameter simulation for Stewart platform design of driving simulator and the inverse kinematics is used as the basis. The dynamic errors between the desired lengths and the actual lengths of legs of the platform are calculated by the simulation model established in MATLAB/Simulink to evaluate the accuracy of the platform. Three common motion scenes including bump, pitch, and yaw are designed for experiment analysis. The dynamic errors are found to increase with the increase of the movement frequency, amplitude, and rotation angle of different motion scenes, and a nearly linear correlation is found between the errors and the movement parameters. The results of the study provide valuable references for the design of physical model in a driving simulator.

Keywords Stewart platform · Inverse kinematics · Movement parameter · Dynamic error · Driving simulator

Introduction

Driving simulators are widely used in the new vehicle deployment, driver training as well as a variety of scientific research. The performance of the simulator's requirement becomes more realistic with the consistent development and applications of

G. Hu · X. Li · X. Yan (✉)
MOE Key Laboratory for Urban Transportation Complex Systems Theory and Technology, Beijing Jiaotong University, Beijing 100044, China
e-mail: xdyan@bjtu.edu.cn

G. Hu
e-mail: 14120832@bjtu.edu.cn

X. Li
e-mail: 13114249@bjtu.edu.cn

driving simulators in recent years. High-level driving simulators provide relative real driving feeling by virtual scene building, and the motion platform enables the driving simulator to motion according to the scene.

A Stewart platform is a type of parallel robot that has six prismatic actuators, commonly hydraulic jacks or electric actuators, attached in pairs to three positions on the platform's baseplate, crossing over to three mounting points on a top plate [1]. A Stewart platform enables devices placed on the top plate to move in the six degrees of freedom which include three linear movements x, y, z (lateral, longitudinal, and vertical) and the three rotations pitch, roll, and yaw [2].

As a typical motion platform of vehicle driving simulator, the Stewart platform has been firstly used in 1965, when Stewart introduced it for training the helicopter pilots. Due to the advantages of high load capacity, high precision, and good repeatability in the design configurations, the Stewart platform has now been widely used in a wide range including researches and developments of automobiles, aircrafts, and bikes [3].

Most of the early Stewart platforms have used hydraulic system as the driving mechanism. As the electric servo drive system and serial servo control technologies have been widely applied in industries, their advantages in low-cost and high-efficiency make the electric servo drive systems gradually replace the traditional hydraulic drive systems in many motion simulation applications, including the Stewart platforms [4].

The typical control strategy of the Stewart platform is realized by the forward and inverse kinematics. Given the specific position and attitude of the platform, the lengths of the six legs in the platform are computed individually. Based on the forward and inverse kinematics, the dynamics model can be established in simulation software, which has a high similarity with the dynamic performance of the real Stewart platform [5].

This study aims to establish a Stewart platform simulation model which is used in the driving simulator in MATLAB/Simulink with the Stewart platform inverse kinematics, and analyze the relationship between the dynamic errors and the motion parameters in different simulation scenes including bump, pitch, and yaw which are the most common motion types in presently existed driving simulators.

Research Method

In the Stewart platform design for a driving simulator, the dynamic errors between the desired lengths of the six legs of the platform and the actual lengths of the legs are normally used to assess the movement accuracy of the platform. In this study, given a specific position and attitude of the platform, the lengths of the legs could be obtained through the established mathematic model based on the inverse kinematics. The simulation model of the platform was realized in MATLAB/Simulink, consisting three parts which were the leg trajectory, the PID (proportion, integral, and derivative) controller, and the mechanical model. Based on the mathematic
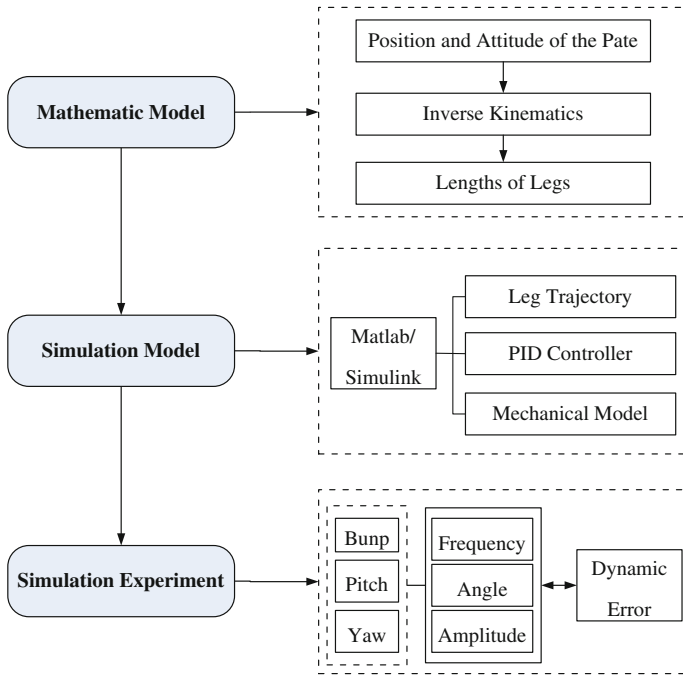
Fig. 1 Research method framework

model and the simulation model, three common motion scenes including bump, pitch, and yaw were chose in the experiment analysis to compare the relationship between the dynamic error and the motion parameters. The research method framework is shown in Fig. 1.

Inverse Kinematics of the Stewart Platform

The simplified Stewart platform with six degrees of freedom is illustrated in Fig. 2. It is composed of a fixed base plate, a movable top plate, and six linear actuators, each of which is connect to both the top plate and the base plate. As the base plate is fixed on the ground, the shortening and extending of the actuator lengths enables the motion of top plate. A coordinate frame $\mathbf{B_i}$ ($i = 1, …, 6$) is applied on the base plate, where O_B is the center and X, Y, Z axes of $\mathbf{B}$ are assumed to be the longitudinal, lateral, and perpendicular directions, respectively. Similarly, a coordinate frame $\mathbf{P_i}$ ($i = 1, …, 6$) is arranged on the top plate, where O_p is the center with X, Y, Z axes as the longitudinal, lateral, and perpendicular directions, respectively. The length of each leg is denoted by l_i ($i = 1, …, 6$).

Fig. 2 Stewart platform coordinate frames
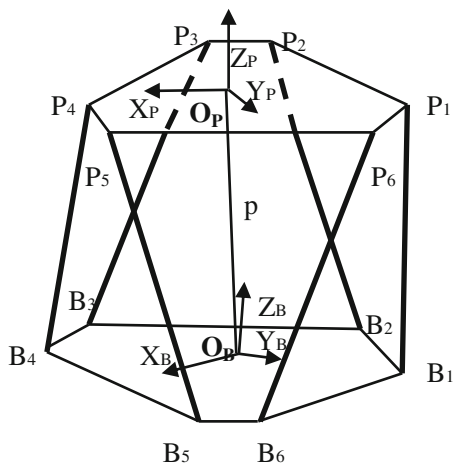
The inverse kinematics for the Stewart platform is used to determine the leg lengths of the top plate **P** for a given position and attitude with respect to the fixed base **B**. Thus, the movement of the moving top plate can be described by the position vector p of the centroid and the rotation matrix R_P^B. Using Euler angle representation, the rotation matrix R_P^B is listed as follows:

$$R_P^B = \begin{bmatrix} c\gamma c\beta & -s\gamma c\alpha + c\gamma s\beta s\alpha & s\gamma s\alpha + c\gamma s\beta c\alpha \\ s\gamma c\beta & c\gamma c\alpha + s\gamma s\beta s\alpha & -c\gamma s\alpha + s\gamma s\beta c\alpha \\ -s\beta & c\beta s\alpha & c\beta c\alpha \end{bmatrix} \tag{1}$$

In which, $c\alpha$, $c\beta$, $c\gamma$ represent $\cos \alpha$, $\cos \beta$, $\cos \gamma$; $s\alpha$, $s\beta$, $s\gamma$ represent $\sin \alpha$, $\sin \beta$, $\sin \gamma$; angle α, β, and γ represent the orientation of frame **P** with respect to frame **B** in X, Y, and Z directions, respectively.

The desired leg position of l_i is shown in Eq. 2, and the length of each leg is shown in Eq. 3.

$$\bar{l}_i = \overline{P_i B_i} = p + R_P^B \cdot b_i^P - a_i^B \tag{2}$$

$$|l_i| = \sqrt{[p + R_P^B \cdot b_i^P - a_i^B]^{\mathrm{T}}} \cdot \sqrt{p + R_P^B \cdot b_i^P - a_i^B} \tag{3}$$

In which, b_i^P represents the vector of joint point i between the leg i and the plate P; and the ith represents the vector of joint point i between the leg i and the plate B, $i = 1, \ldots, 6$.

The lengths of six legs of the platform will determine the position and attitude of the top plate, so inversely, given any position and attitude of the top plate within its working space, the lengths of six legs can be calculated with Eq. 3.

Program Simulation in MATLAB/Simulink

Based on the inverse kinematics model, a simulation program was developed using MATLAB/SimMechanics, which consists of three parts of leg trajectory, PID controller, and mechanical model, as is show in Fig. 3.

(1) Leg trajectory

The leg trajectory controller is used to keep the top plate move as the instruction. As the desired attitude and position of the top plate are achieved by controlling the force implemented on the driving legs to change the lengths of them, the actual position of the top platform is determined by the movement of the legs. In addition, the position and attitude of top plate is much more difficult to measure than the lengths of the legs. Thus, by using the forward-inverse kinematics solutions between them, the desired position and attitude of top plate are transformed to the legs' position to achieve the targeted movement of the top plate.

(2) PID low-level controller

The PID controller aims to decrease the error between desired leg position and actual leg position by building a closed-loop system with the mathematics models (as shown in Fig. 4). Thus, each leg has an independent PID controller. There are three input parameters including the actual leg position, the desired leg position, and the movement velocity. In the PID controller, the desired leg position is generated by the leg trajectory system, and the actual leg position and velocity are the output of the mechanical model. By comparing the error between the desired and the actual leg position, the controller generates appropriate force to be input into the mechanical model to reduce the error.
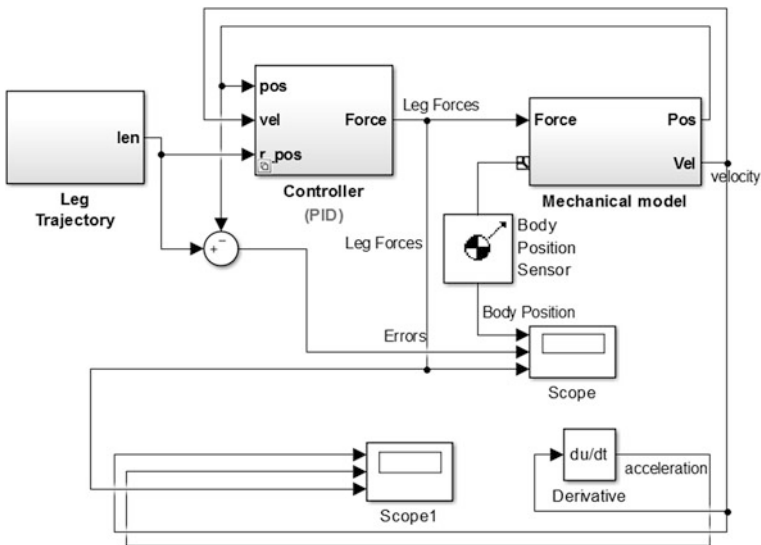
Fig. 3 Simulation program
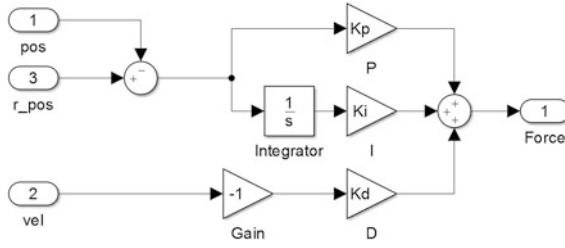
Fig. 4 PID controller
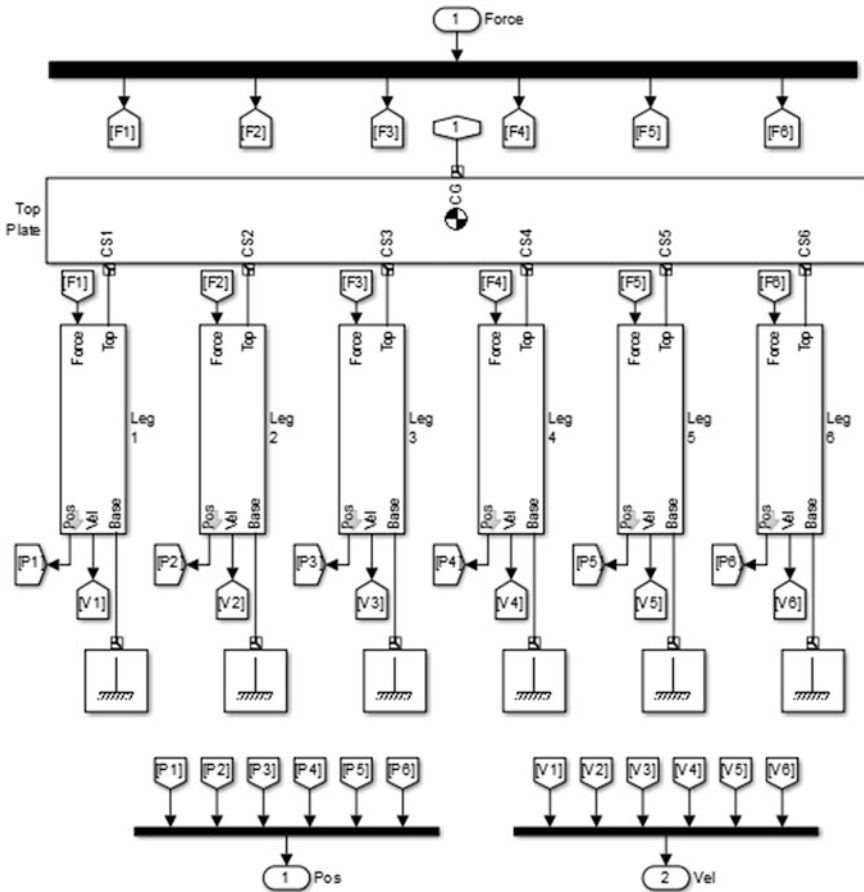
Fig. 5 Mechanical model of Stewart platform

(3) Mechanical model

The mechanical model of Stewart platform is built in SimMechanics (as shown in Fig. 5). As a feature in Simulink, the SimMechanics is a block diagram

model environment which allows user to model and simulate mechanical systems by specifying the bodies with mass properties, possible motions, and kinematic constraints [3]. The mechanical model mainly consists of a top plate, a base plate, and six legs connecting the top plate and the base plate. The ground block acts as the fixed base plate in the model. Each leg subsystem is composed of two bodies, two universal joints, and one adjustable cylindrical joint. The upper body of each leg connects to the top plate using a universal joint, and the lower body connects to the base plate using another universal joint. The overall system has six degrees of freedom.

Simulation Experiment

The simulation experiment was conducted to assess the movement accuracy of the platform established in the study by analyzing the dynamic errors in different motion scenes. Assuming that the Stewart platform was used for a driving simulator design, three common movements including bump, pitch, and yaw were involved in the simulation experiment. In the bump scene, the top plate repeats the vertical reciprocating motion along Z-axis, which simulates the vehicle driving in a bumpy situation. In the pitch scene, the top plate keeps revolving around Y-axis, simulating the uphill and downhill driving of a vehicle. In the yaw scene, the top plate keeps revolving around the Z-axis, which simulates the steering wheel turning behavior of a vehicle. The movement trajectories of the top plate in the three motion scenes were all assumed to follow sine function. As the movement range of the platform is limited, the other motion scenes are not considered in the study. Some main information about volume and mass of the platform in the experiment is listed in Table 1, and the real-time movement of the platform is shown in Fig. 6. According to the input movement parameters such as frequency, amplitude, and angle rotation within a reasonable range in different motion scenes, the dynamic error output of the platform is shown in Fig. 7.

Table 1 Main information of platform

Related parameters	Value
The mass of load on the top plate	500 kg
The mass of upper leg	9.71 kg
The mass of lower leg	5.46 kg
The base plate radius	1.5 m
The top plate radius	1 m
The height of top plate	1 m
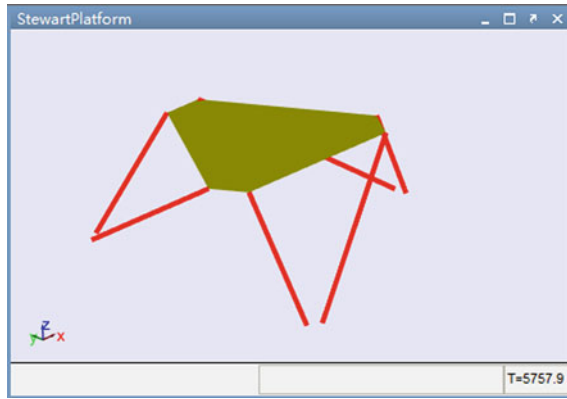
Fig. 6 3D graph of Stewart platform
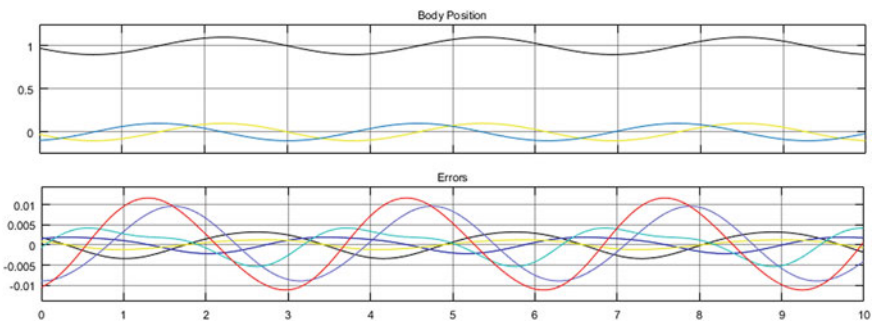
Fig. 7 Dynamic error output of Stewart platform

Table 2 Errors in different movement frequency and amplitude conditions in bump scene (m)	f (Hz)	A (m)				
		0.01	0.02	0.03	0.04	0.05
	2	0.0022	0.0042	0.006	0.008	0.01
	4	0.0044	0.009	0.0134	0.0175	0.022
	6	0.0072	0.014	0.021	0.028	0.036
	8	0.008	0.017	0.026	0.035	0.045
	10	0.0095	0.018	0.028	0.036	0.046

Simulation Results and Discussion

Simulation of Bump Scene

As the movement trajectory of the top plate follows sine function ($z = A \sin ft$), the movement amplitude A and frequency f are the control variables for testing the dynamic errors. Based on the simulation model, the errors between the desired lengths of legs and the actual lengths of legs in bump motion with different amplitudes and frequencies were calculated and listed in Table 2.
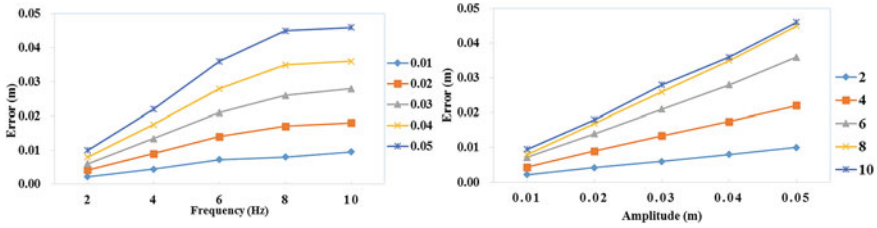
Fig. 8 Relationship between errors, movement frequency, and amplitude in bump scene

Table 3 Errors in different movement frequency and rotation angle conditions in pitch scene (m)

f (Hz)	β (°)				
	5	10	15	20	25
1	0.0009	0.0017	0.0025	0.0034	0.0042
2	0.0018	0.0034	0.0051	0.0067	0.0084
3	0.0025	0.005	0.0075	0.010	0.0125
4	0.0032	0.0066	0.010	0.013	0.016
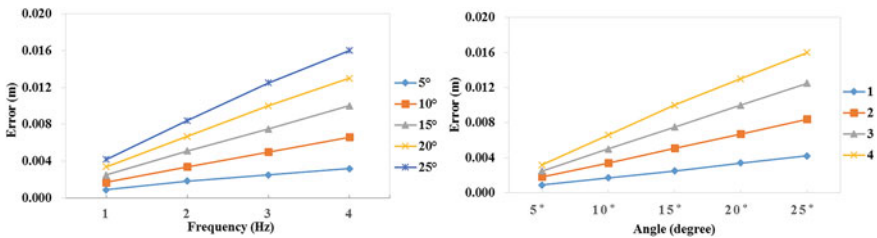
Fig. 9 Relationship between errors, movement frequency, and rotation angle in pitch scene

As shown in Fig. 8, the error increased with the increase of frequency and amplitude. It also shows that when the frequency was large enough, the error differences between different frequency conditions were not obvious, which indicates that the error would keep constant for a certain amplitude when the frequency was over 8 Hz. Consequently, high frequency and large amplitude of bump motion can cause large errors while in the driving scene, the bump motions with high frequency and large amplitudes usually come from driving on rough roads at high speed.

Simulation of Pitch Scene

As the top plate of the platform keeps revolving movement around *Y*-axis in pitch motion, the input variables in the sine function that the motion trajectory follows are the movement frequency *f* and the rotation angle *β*. The dynamic errors in different frequency and rotation angle conditions are listed in Table 3, and the profiles of change are shown in Fig. 9. The error increases with the increase of frequency and

rotation angle, and it is obvious that the increase presents an approximately linear tendency. Combined with the driving scene, the results indicate that large and frequent uphill or downhill slopes condition will increase the error of the platform.

Simulation of Yaw Scene

In the yaw scene, the top plate keeps revolving around the Z-axis, and the movement frequency f and rotation angle Γ of the sine function are the control variables. Table 4 displays the dynamic errors in different frequency and rotation angle conditions, and Fig. 10 shows the error change with the frequency and rotation angle. Similar to the tendency in pitch scene, the error increases with the increase of frequency and rotation angle, and there is a nearly linear correlation between the error and the control variables. Thus, in addition to the vertical rotation, wide and frequent horizontal rotation can also increase the platform error, for instance, the fast and significant wheel steering on curve road.

In the design of the Stewart platform, the error between the desired position and actual position of each leg is the key indicator to evaluate the performance of the platform [6]. The three groups of simulation experiments show a common tendency that the errors increase with the movement frequency, amplitude, or rotation angle. Similarly, Lou and Tseng found that as the frequency of sinusoidal movement command increases, the motion of the platform becomes more awkward and deviated from the commands [3]. The changes of errors in pitch and yaw scenes are within 0.02 m according to the established model, but it needs to be mentioned that the error exhibited in the present study is the maximum value within a motion

Table 4 Errors in different movement frequency and rotation angle conditions in yaw scene (m)

f (Hz)	Γ (°)			
	5	10	15	20
0.5	0.00050	0.0010	0.0014	0.0020
1	0.00085	0.0019	0.0027	0.0038
1.5	0.00150	0.0030	0.0043	0.0057
2	0.00185	0.0037	0.0055	0.0074
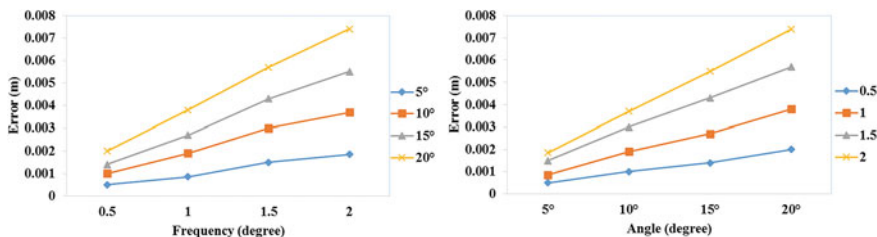
Fig. 10 Relationship between errors, movement frequency, and rotation angle in yaw scene

period while in most previous studies, only the minimum error was involved [7, 8]. The error in bump scene is relatively larger with a maximum value of 0.046 m compared to the other two motions but this is probably because of the high movement frequency, and once the frequency reaches 8 Hz, the error keeps constant approximately. In addition, the nearly linear correlation found between the dynamic errors and the movement parameters in the study provides potential likelihood to calibrate the outputs of the platform. For the three scenes considered in the study, the experiment results indicate that to limit the error within acceptable range, the implementation of a driving simulator in virtual road scenes with respect to the road roughness, the slope of uphill and downhill, the curvature of curve road, as well as the speed limit could be regarded as adjustable factors.

Conclusion

This study investigates the application of Stewart platform in a driving simulator by evaluating the dynamic errors in different motion scenes. The mathematical model of the platform is based on the inverse kinematics, and the simulation model is established by using MATLAB/Simulink. Three motion scenes including bump, pitch, and yaw are involved in experiment simulation to analyze the relationship between dynamic errors and the movement parameters. The results indicate that the dynamic errors increase with the increase of the movement frequency, amplitude, and rotation angle of different motion scenes, and there is a nearly linear correlation between the errors and the movement parameters. The study sheds some lights on the driving simulator design in various road scenarios and driving conditions to achieve high simulation accuracy.

References

1. Lazard, D., and J. -P. Merlet. 1994. The (true) Stewart platform has 12 configurations. In *Proceedings of the 1994 IEEE International Conference on Robotics and Automation*, p. 2160.
2. Tu, Kuo-Yang, Tung-Chung Wu, and Tsu-Tian Lee. 2004. A study of Stewart platform specifications for motion cueing systems. In *2004 IEEE International Conference on Systems, Man and Cybernetics*, vol. 4, pp. 3950–3955.
3. Ravichandran, K. 2010. Driving simulator for tracked vehicles using Stewart platform manipulator. In *2010 International Conference on Emerging Trends in Robotics and Communication Technologies (INTERACT)*, Chennai, 2010, pp. 245–249.
4. Lou, J. H., and S. P. Tseng. 2014. Developing a real-time serial servo motion control system for electric stewart platform. In *2014 International Conference on Advanced Robotics and Intelligent Systems (ARIS)*, Taipei, 2014, pp. 66–71.

5. Qian, Zhou, and Zhang Hui. 2003. Simulation and experimental analysis of the Stewart parallel mechanism for vibration control. In *IEEE International Conference on Systems, Man and Cybernetics, 2003*, vol. 4, pp. 3548–3552.
6. Du, Yihao, Ping Xie, and Bin Liu. 2010. Error analysis in parallel robot using Edgeworth series and information entropy. In *2010 International Conference on Computer Design and Applications (ICCDA)*, Qinhuangdao, pp. V2-350–V2-354.
7. Dastgerdi, H. R., and M. Keshmiri. 2010. Design of length measuring system for Stewart platform using new forward kinematics solution. In *2010 11th International Conference on Control Automation Robotics & Vision (ICARCV)*, Singapore, pp. 2339–2344.
8. Dai, X., Q. Huang, J. Han, and H. Li. 2009. Accuracy synthesis of Stewart platform used in testing system for spacecraft docking mechanism. In *2009 International Conference on Measuring Technology and Mechatronics Automation*, Zhangjiajie, Hunan, pp. 7–10.

Grid-Based Framework for Railway Track Health Evaluation

Lei Bai, Qiuyan Zhang, Rengkui Liu, Futian Wang and Xiaoqi Song

Abstract It is imperative for railway managers to have an accurate grasp of the track health to optimize the scheduling of maintenance and replacement, and reasonably allocate the maintenance and replacement resources. In this paper, based on the gridding management theory, a railway track is divided into several adjacent segments of the same specific length. A condition evaluation index system and a track grid health index are established for the divided segments. The track grid health evaluation framework is proposed for calculating the track grid health index. The proposed framework was verified with the track performance data from Lanzhou–Xinjiang railway line. Our evaluation is demonstrated that the proposed track grid health index outperforms the track quality index and track quality rank, which are the traditional condition indicators for evaluating track quality in China's railway management.

Keywords Railway transportation · Track health · Gridding management · Clustering learning

Introduction

Railway track, as a key railway facility, is the foundation for train operation. It guides train running, provides a surface for wheels, and spreads the load of wheels to the underlying subgrade [1]. Track health significantly influences traveling safety of trains and passenger comfort.

L. Bai (✉) · F. Wang
State Key Laboratory of Rail Traffic Control and Safety, Beijing Jiaotong University,
Beijing 100044, China
e-mail: blbailei@gmail.com

Q. Zhang · R. Liu
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China

X. Song
Lanzhou Bureau of China Railways, Lanzhou 730000, China

As trains are developed into higher speeds with higher axle loads, maintenance and replacement (M&R) activities are increasingly required and costly to maintain the railway track in a good condition. An accurate grasp of track health is of great significance to the optimal scheduling of M&R, optimal allocation of M&R resources, and cost optimization. It is a big challenge faced by the railway managers to make trade-off decisions to map various track condition indicators onto the comprehensive track health evaluation, adapting to the trend of centralized M&R.

Ballasted track [1] is the structure consisting of rails, sleeper, fasteners, and ballast. Slab ballastless track [2] is the structure consisting of rails, fasteners, slab tracks, cement asphalt mortar, and baseplates. Depending on the above definition, railway track is the structure consisting of multiple facilities. Domestic and foreign research mainly focuses on the condition assessment and deterioration law of a kind of facility among the railway track [3, 4], while seldom research is about the comprehensive evaluation of track health [5]. The evaluation of track irregularities and deterioration law is mainly addressed [6–9], but track irregularity is the only one aspect of measurement for track health. Based on RAMS (Reliability, Availability, Maintainabilty, Safety) [10], Kang et al. [9] proposed a new comprehensive evaluation method for track irregularities. Sadeghi and Askarinejad [5] developed a new track structure assessment technique, using engineering experiences and expert judgments. Track structural index is evaluated using the distresses of track components, such as rails, sleeper, fasteners, and ballast.

In other infrastructure management field, scoring systems are employed for comprehensive facility condition evaluation. The relative weight of attributes is obtained, using different methodologies, such as the Delphi technique, analytic hierarchy process (AHP), fuzzy logic, and expert system method [11–17]. These methodologies are based on engineering experiences and expert judgments.

Haas et al. [15] employed the Delphi technique for the pavement condition assessment. Liang et al. [16] adopted a fuzzy synthetic evaluation method to evaluate the condition of existing reinforced concrete bridges. Sun and Gu [11] developed a new approach for evaluating pavement condition, using AHP and fuzzy logic theory. Koduru et al. [12] employed fuzzy logic and expert system techniques for pavement condition evaluation. Bianchini [14] proposed a fuzzy inference model to obtain pavement condition ratio (PCR) for pavement condition assessments. Zhang et al. [13] developed a fuzzy AHP synthetic evaluation model for the health rating scores of shield tunnels, using different types of measurement data. Zayed et al. [17] proposed an integrated approach of the analytical hierarchy process with the multi-attribute utility theory for calculating condition-rating index, to evaluate wastewater treatment plants and prioritize the rehabilitation activities.

In this paper, a railway track is divided into several adjacent segments of the same specific length. These segments are called track grids and are used as research objects. A condition evaluation index system of track grids is established. Employing PCA and hybrid hierarchical k-means clustering model, the railway track grid health evaluation framework (TGHEF) is developed for calculating track grid health index (TGHI) to help managers comprehensively grasp the track health.

Our paper is different from other researchers' approaches in two standpoints.

- A railway track is divided into several adjacent track grids, and track health can be accurately grasped in a smaller spatial scale. A condition evaluation index system of track grids is established to cover the shortage of existing condition assessment system of railway track.
- The railway track grid health evaluation framework is developed. The data-driven framework can decrease managers' subjective influence on evaluating the health of track grids.

The structure of the remainder of this paper is as follows: The definition and division of track grids are covered in Section "Definition and Division of Track Grids". And Section "Condition Evalutaion of Track Grids" proposes a condition evaluation index system of track grids and track grid health index. Railway track grid health evaluation framework for calculating TGHI is discussed in Section "Railway Track Grid Health Evaluation Framework". Section "Empirical Analysis" utilizes measurement data to demonstrate the validation of the proposed framework. Our conclusions are presented in Section "Conclusions".

Definition and Division of Track Grids

The track line exhibits a continuous, belt-like distribution. The track line is divided into adjacent segments of 200 m length, which are called track grids. The health condition of rails, sleeper, ballast and fasteners is evaluated within the specific length. Take the ballasted track as an example, a track grid is the structure consisting of the rails, sleeper, ballast, and fasteners within the specific length. The track health condition is determined by the condition of track components (such as rails, sleeper, ballast, and fasteners) within the specific length. And it is the overall performance of multiple track components.

The main reasons to divide the track line into track grids in units of 200 m can be summed up in three points.

- Heterogeneity is one of the characteristics of railway infrastructure condition [6, 18]. It means that infrastructure of the same type has different degradation rules at different spatial locations. Due to the heterogeneity, it is appropriate for the continuous linear rails and ballast to be divided into segments to evaluate the health condition and the health deterioration law.
- The degradation of railway infrastructure is also characterized by linkages [18]. Specifically, the degradation processes interact among different infrastructures with the same mileage (or within the same mileage range). Failure or fault in one infrastructure would often lead to failures or faults in associated infrastructure. For example, the distresses of fasteners have an extremely high probability of causing track irregularities. The railway managers should comprehensively evaluate health condition of multiple track components.
- The M&R capability and management level of the China's railway department.

Condition Evalutaion of Track Grids

A Condition Evaluation Index System of Track Grids

A condition evaluation index measures the performance of railway track. It guides the goal of railway managers' maintenance decision-making and gives directions. A condition evaluation index affects the maintenance of persons' objectives and behaviors. Inappropriate condition assessment indicators will result in unreasonable M&R activities.

In the China's railway management, the shortages of the existing condition assessment index system of railway track [19] are as follows:

- These existing indicators are designed without consideration for heterogeneity [18]. It is difficult to precisely describe and master linear continuous track components condition, using these indicators.
- These existing indicators are designed without consideration for linkages [18]. They mainly measure the performance of track components, such as rails and sleepers, while lacking comprehensive condition indexes for evaluating track health.
- These existing indicators lack the quantitative evaluation of condition change trends, such as *average failure rate* and *failure concentration rate*.
- These existing indicators lack the quantitative evaluation of safety.
- These existing indicators lack the quantitative evaluation of aesthetics.

To cover the above shortages, a condition evaluation index system is established and used to measure the performance track grids, from the four aspects such as functional performance, structural integrity, safety, and aesthetics. Take ballasted track as an example to illustrate the condition evaluation index system of track grids, as shown in Table 1.

Track Grid Health Index

Different condition indexes, as shown in Table 1, have different measures on the performance of track grids. Track grid health index (*TGHI*) is, therefore, proposed to comprehensively evaluate track grid health in this paper. *TGHI* is a discrete quantitative index of track health. Different *TGHI* values represent different conditions of track grids. *TGHI* is a function of condition indicators shown in Table 1, as follows:

$$TGHI = \underset{i \in M}{f} (q_i) \tag{1}$$

wherein $q_i (i \in \{1, 2, \ldots, M\})$ represents the condition index as shown in Table 1. M expresses the total number of condition indicators of track grids. f represents the

Table 1 A condition evaluation index system of track grids (ballasted track)

Class	Indicators
Functional performance	*Track regularity index* [20] (such as surface, alignment, gauge, cross-level, and twist) *Track quality index* (*TQI*) [19, 21], *Change rate of TQI, Track quality rank* (*TQR*), *T* [8, 21]
	Track dynamic stiffness
Structural integrity	*Rail defects*
	Ballast defects
	Sleeper and fasteners defects
Safety	*Average failure rate*
	Change rate of failure quantity
	Mean time between failures
	Mean time between M&R
	Hazard rate [8]
	Service life
	Failure repetition rate
	Failure concentration rate
Aesthetics	*Cleanliness rate of rail surface*
	Dirty rate of ballast
	Dirty rate of sleeper
	Dirty rate of fastener

mapping relationship between q_i and *TGHI*. How to obtain *TGHI* of track grids will be addressed in Section "Railway Track Grid Health Evaluation Framework".

Railway Track Grid Health Evaluation Framework

Core Components

The core components of the railway track grid health evaluation framework (RTGHEF) are as follows:

- The railway line is divided into track grids, and the track health can be accurately grasped in a smaller spatial scale by managers.
- Employing PCA model [22], the condition indicators (or variables), as shown in Table 1, are transformed into a new small set of condition variables, and the dimensionality of the multivariate data measuring track condition is reduced.
- Based on new condition variables, the *TGHI* of track grids is obtained, employing the hybrid hierarchical k-means clustering model.

PCA Model

Q represents a set of condition indicators (or variables) of track grids. $Q = (q_1, q_2, \ldots, q_i, \ldots, q_M)^T$. μ expresses the mean of Q, $\mu = E(Q)$. σ expresses the variance of Q, $\sigma = Var(Q)$. Z represents principal components, $Z = (z_1, z_2, \ldots, z_i, \ldots, z_M)^T$. The new condition variable (or principal component) Z corresponds to a linear combination of the condition variable Q, as shown in Eq. (2). A represents the linear combination coefficients, $A = (a_1, a_2, \ldots, a_i, \ldots, a_M)^T$ $a_i = (p_{1i}, p_{2i}, \ldots, p_{ji}, \ldots, p_{Mi})^T$. p_{ji} represents linear combination coefficients between z_i and q_j.

$$\begin{cases} z_1 = a_1^T Q \\ z_2 = a_2^T Q \\ \vdots \\ z_I = a_I^T Q \end{cases} \tag{2}$$

Based on Eq. (2), the variance of z_i, $Var(z_i) = a_i^T \sigma a_i$ $(i, j = 1, 2, \cdots, M)$. The core of PCA is to search the appropriate linear combination coefficient A to make the variance of Z, $Var(Z)$ maximum. That is to say, the optimal linear combination coefficients $\widehat{A}$ could be determined by the solution of the constrained optimization problems as shown below.

$$\max \widehat{A} \sigma A$$
$$s.t. \widehat{A} A = 1 \tag{3}$$

This optimization problem is equivalent to calculating the eigenvectors corresponding to maximum eigenvalues of σ. It is assumed that $\lambda_1, \lambda_2, \cdots, \lambda_M$ represent the eigenvalues of σ. $\widehat{a}_1, \widehat{a}_2, \ldots, \widehat{a}_I$ represent the corresponding eigenvectors of σ. So $\widehat{A} = (\widehat{a}_1, \widehat{a}_2, \ldots, \widehat{a}_I)^T$ represents the optimal linear combination coefficients. The percentage of variances of the principal component z_i is $\lambda_i / \sum_{j=1}^{I} \lambda_j$. The accumulated percentage of variances of the first k principal components is $\sum_{i=1}^{k} \lambda_i / \sum_{j=1}^{I} \lambda_j$. The correlation of the condition indicator q_j with the principal component z_i is $\sqrt{\lambda_i} \widehat{p}_{ji}$.

Hybrid Hierarchical k-Means Clustering Model

It is assumed that the first n principal components are selected as new condition variables to measure the performance of track grids. The total number of track grids

is D, and the discrete value of the TGHI is S. The D track grids are splitting into a set of S clusters, using the hybrid hierarchical k-means clustering model. The hybrid hierarchical k-means clustering model is developed, combining k-means [23] and agglomerative hierarchical clustering. The Euclidean distance is used to measure the dissimilarity between track grids and clusters. An outline of the hybrid hierarchical k-means clustering algorithm used to estimate $TGHI$ of track grids is as follows:

Step 0: Using agglomerative hierarchical clustering, track grids are cut into S clusters based on new condition variables $z_1, z_2, \cdots, z_n$.

Step 1: Compute each center $C_1, \ldots, C_s, \ldots, C_S$ of S clusters, refer to Eq. (4). G_{is} represents the ith track grid among the cluster C_s. t_s represents the total number of track grids among the cluster C_s. $Z_{is} = \left(z_1^{is}, z_2^{is}, \ldots, z_n^{is}\right)^T$ represents the condition variables of the track grid G_{is}.

$$C_s = \frac{1}{t_s} \sum_{i \in t_s} Z_{is} \tag{4}$$

Step 2: Compute the dissimilarity between each track grid and each center $C_1, \ldots, C_s, \ldots, C_S$ of S clusters. Assign each track grid to their most similar cluster, refer to Eq. (5). Z_i represents the condition variable of the track grid G_i. L_s^i represents that G_i is assigned to C_s.

$$L_s^i = \arg \min_{i \in D, s \in S} \|Z_i - C_s\| \tag{5}$$

Step 3: Return to Step 1 until the S clusters formed stop changing.

Step 4: Return to the current clustering.

Empirical Analysis

Overview

The Lanzhou–Xinjiang railway line, which runs between Beijing and Urumqi, is a ballasted track located in an important position in the new Eurasian Continental Bridge. A total of 3683 track grids, ranging from 548 to 985.6 km in UP and DOWN lines, are used as research objects. The framework is verified by the track performance data accumulated in March 2016. Eight indexes $Q = (q_1, q_2, \ldots, q_8)^T$ are prepared for the 3683 track grids for evaluation. Wherein, the index q_1 expresses

TQI, the index q_2 expresses *change rate of TQI*, the index q_3 expresses *track quality rank (TQR)*, the index q_4 expresses *sleeper and fasteners defects*, the index q_5 expresses *average failure rate*, the index q_6 expresses *change rate of failure quantity*, the index q_7 expresses *failure concentration rate*, and the index q_8 expresses *hazard rate*. The size of the corresponding historical performance dataset is 893.

Based on the opinions of experts in Lanzhou Railway Bureau, the *TQR* is proposed for the evaluation of the track quality and applied at present. The index q_3 is determined by the value of *TQI* and the total number of track geometry defects. The index q_3 is categorized into four ranks, and the track quality deteriorates from condition rank I to IV. That is to say, the condition rank I represents the best track quality, while the condition rank IV represents the worst track quality. The condition rank classification standard for *TQR* in Lanzhou–Xinjiang railway is shown in Table 2.

Analysis of Results

The railway track grid health evaluation framework is built and *TGHI* is obtained, using R language [24, 25]. z_1 represents the first principal component of the condition indicators $Q = (q_1, q_2, \ldots, q_8)$. z_2 represents the second principal component. The percentage of variances of z_1 is 45.0%. The percentage of variances of z_2 is 17.6%. The accumulated percentage of variances of the first two principal components is 62.6%. Because the accumulated percentage exceeds 60%, first two principal components, z_1, z_2, are selected as new condition variables to measure the performance of track grids.

The correlation of the condition indicator Q with the first two principal components, z_1 and z_2, is shown in Table 3. The condition indicator Q on the first two principal components map is shown in Fig. 1. The abscissa value of the arrow is the correlation of q_i with z_1. The ordinate value of the arrow is the correlation of q_i with z_2. The closer a condition indicator q_i is to the circle, the better its representation by the first two principal components.

Table 2 Condition rank of *TQR*

TQR	Determining standard
I	$TQI < 6.5$, $NTGD_{tgt} < 8$, and $NTGD_{pmi} < 6$
II	$TQI \geq 6.5$; or $NTGD_{tgt} \geq 8$; or $NTGD_{pmi} \geq 6$;
III	$TQI \geq 6.5$, $NTGD_{tgt} \geq 8$; or $TQI \geq 6.5$, $NTGD_{pmi} \geq 6$; or $NTGD_{tgt} \geq 8$, $NTGD_{pmi} \geq 6$; or $NTGD_{tgt} \geq 20$
IV	$TQI \geq 6.5$, $NTGD_{tgt} \geq 8$, and $NTGD_{pmi} \geq 6$; or $NTGD_{tgt} \geq 30$

Note The number of track geometry defects from the track geometry trolley is referred to as simply $NTGD_{tgt}$. The number of track geometry defects from the portable measuring instrument is referred to as simply $NTGD_{pmi}$

Table 3 Correlation of the condition indicators with principal components

Correlation	z_1	z_2
q_1	0.496	0.620
q_2	−0.430	0.648
q_3	0.861	−0.018
q_4	0.080	0.617
q_5	0.948	0.021
q_6	0.005	0.452
q_7	0.948	0.021
q_8	0.792	−0.131
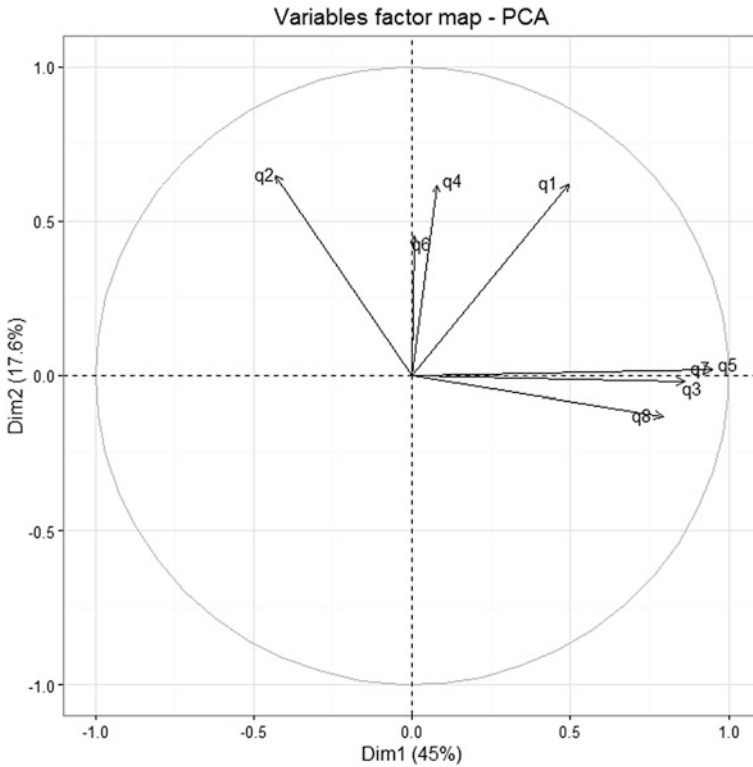
Fig. 1 Condition indicators on the first two principal components map

Table 3 and Fig. 1 show that z_1 is strongly positively correlated with the condition indicators q_1, q_3, q_5, q_7, and q_8, and negatively correlated with the condition indicator q_2. z_1 represents that track grids have higher value of q_1, q_3, q_5, q_7, and q_8, and have lower value of q_2. And z_1 mainly describes the quantity of defects. The strongly positively correlated variables with z_2 are the condition indicators q_1, q_2, q_4, and q_6. z_2 represents that track grids have higher value of q_1, q_2, q_4, and q_6. And z_2 is
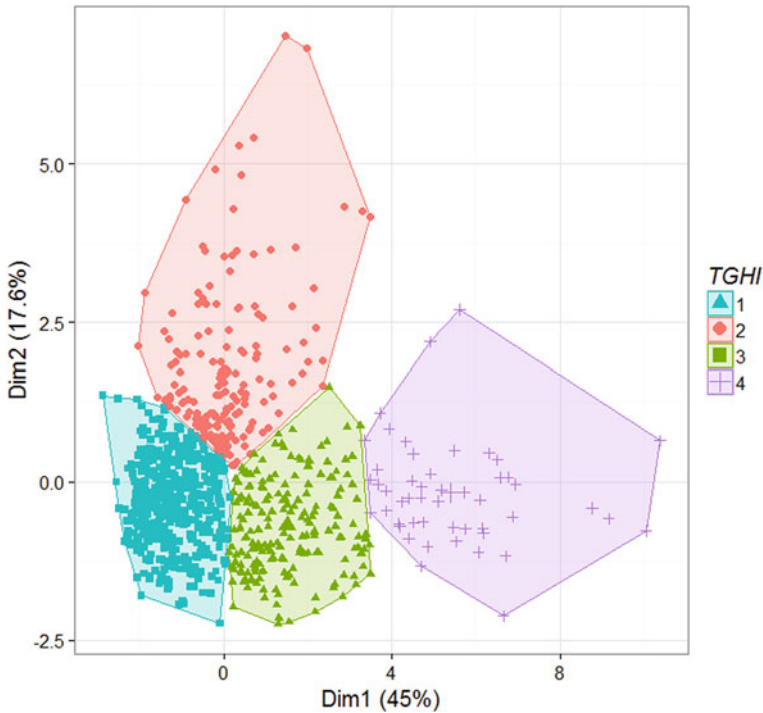
Fig. 2 Clustering map of track grids

mainly used to describe the change trend of defects and the quantity of sleeper and fasteners defects.

The clustering results of the track grids, ranging from 548 to 985.6 km in UP and DOWN lines, are shown in Fig. 2, where the horizontal axis represents the principal component z_1 and the vertical axis represents the principal component z_2. These track grids can be categorized into four labels, and corresponding health feature and maintenance activities are listed in Table 4.

According to the health feature of track grids, railway managers need to assign prior maintenance activities for the track grids, *TGHI* of which is 2 or 4. And the inspection interval for the track grids, *TGHI* of which is 2, needs to be shortened.

Evaluation of Results

To verify the *TGHI* of track grids and the proposed RTGHEF framework, we use the *TQI*, which is an important index of track quality in China's railway management [19], to evaluate the health of track grids and to compare the ranking results with *TGHI*. *TQI* is employed to assess the whole section dynamic track irregularity.

Table 4 Clustering results

TGHI	Legend	Health feature			Maintenance activities
		Average failure rate, hazard rate, TQI, TQR	Change rate of TQI, change rate of failure	Sleeper, fasteners defects	
1	Triangular points	Low	Slow	Few	Do noting
2	Circular points	Average	Fast	Many	Assign prior maintenance activities
3	Square points	Average	Slow	Few	Plan maintenance activities and arrang them within a suitable time
4	Cruciform points	High	Slow	Few	Assign prior maintenance activities

Table 5 Top ten worst health track grids

Railway line	Direction	Location (km)	TQI (mm)	TQI ranking	TGBI
Lanzhou–Xinjiang	Down	722.6	9.77	1	2
Lanzhou–Xinjiang	Up	727.6	9.64	2	2
Lanzhou–Xinjiang	Up	723	9.59	3	4
Lanzhou–Xinjiang	Down	654.8	9.58	4	2
Lanzhou–Xinjiang	Down	654	9.19	5	2
Lanzhou–Xinjiang	Down	652.8	8.88	6	2
Lanzhou–Xinjiang	Down	655.4	8.38	7	2
Lanzhou–Xinjiang	Down	722.2	8.23	8	2
Lanzhou–Xinjiang	Up	908.6	8.21	9	3
Lanzhou–Xinjiang	Down	650.2	7.95	10	2

It indicates the dispersion degree of the track irregularity. A large *TQI* implies a poor regularity and more fluctuation in the track sections.

Using the *TQI*, the health condition of track grids is ranked. Based on the ranking results, the top ten worst health track grids are obtained, as shown in Table 5. The corresponding track grids on the first two principal components map are as shown in Fig. 3.

Table 5 and Fig. 3 show that the top ten track grids mainly belong to the track grid set, *TGHI* of which is 2. If managers make M&R decisions based on *TQI*, the track grid set, *TGHI* of which is 2, will be assigned prior maintenance activities. However, the maintenance activities for the track grid set, *TGHI* of which is 4, may not be planned. So the allocation of M&R resource will be inefficient, and it will even lead to railway accidents.

Based on the *TQI*, experts of Lanzhou Railway Bureau propose an improved index, *track quality rank* (*TQR*), to comprehensively evaluate the track quality and prioritize the M&R activities and to further verify the *TGHI* and the proposed
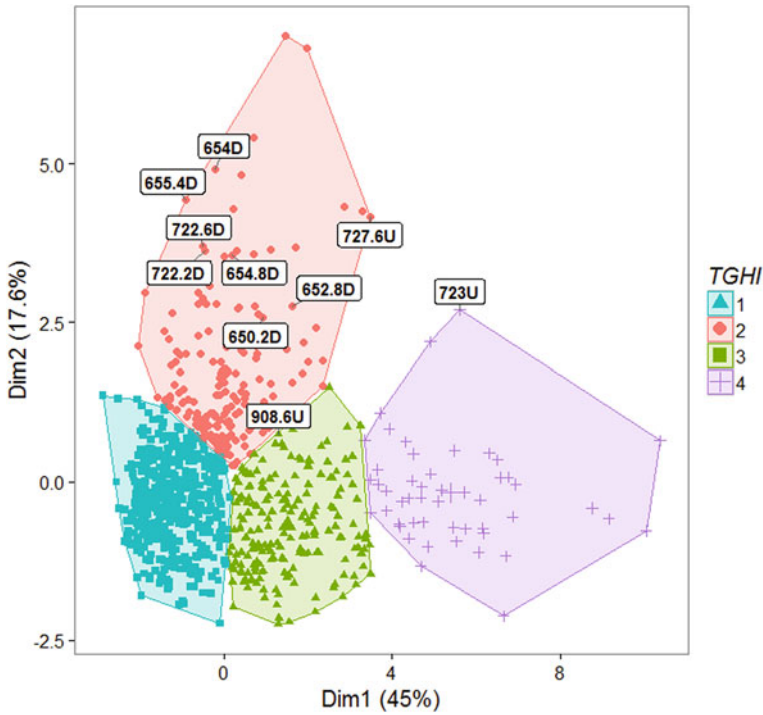
Fig. 3 Top ten worst health track grids on the first two principal components map

Table 6 Confusion matrix for TQR and TGHI

		TGHI				Sum
		1	2	3	4	
TQR	I	392	54	6	0	452
	II	81	108	91	1	281
	III	2	11	96	21	130
	IV	0	0	2	28	30
Sum		475	173	195	50	

RTGHEF framework. We compare the *TGHI* evaluation results with the corresponding *TQR* evaluation results.

A confusion matrix is used mainly to compare the *TQR* and the *TGHI* of track grids. Each matrix element represents the number of track grids, where the *TQR* is the row and the *TGHI* is the column. The main diagonal elements represent the track grids, the *TQR* of which is equal to the *TGHI*.

The confusion matrix for the *TQR* and *TGHI* of track grids is shown in Table 6. A scatter plot of the *TQR* and *TGHI* of track grids is shown in Fig. 4. The horizontal
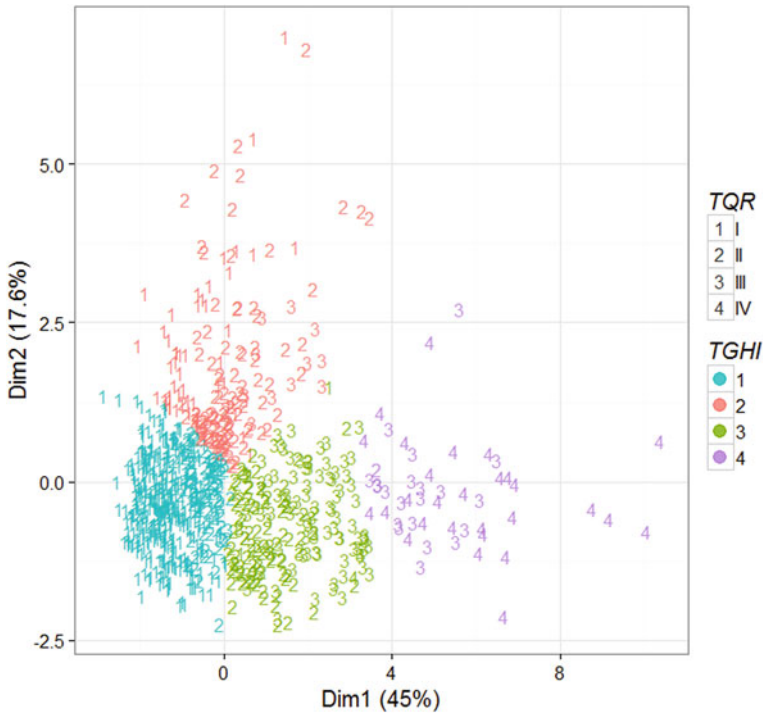
Fig. 4 Scatter plot of the *TQR* and *TGHI*

axis represents the principal component z_1, and the vertical axis represents the principal component z_2. Differences in the *TGHI* of track grids are represented by different colors. Different *TQR* of track grids is represented by different numbers.

Table 6 and Fig. 4 show that track grids, *TQR* of which is I, mainly correspond to track grids, *TGHI* of which is 1. Track grids, *TQR* of which is II, mainly correspond to track grids, *TGHI* of which is 1, 2 or 3. Track grids, *TQR* of which is III, mainly correspond to track grids, *TGHI* of which is 2 or 3. Track grids, *TQR* of which is IV, mainly correspond to track grids, *TGHI* of which is 3 or 4.

Refering Table 2, the *TQR* lacks the quantitative evaluation for change trends of track condition and the number of sleeper and fasteners defects. So track grids, *TGHI* of which is 2, mainly correspond to track grids, *TQR* of which is I, II. These track grids account for 93.5%. Maintenance activities for these track grids may not be arranged if managers make M&R decision based on the *TQR*. The scheduling of M&R and the allocation of M&R resources will be inefficient or irrational.

Conclusions

Based on the gridding management theory, this paper addressed a novel railway track grid health evaluation framework, TGHEF. The track line was divided into several track grids. And railway managers can accurately grasp the track health in a smaller spatial scale. From the four aspects of functional performance, structural integrity, safety, and aesthetics, the evaluation system of track grids was established to cover the shortage of existing condition assessment system of railway track. On this basis, track grid health index was calculated, employing PCA and hybrid hierarchical k-means clustering model. Based on the calculated results and the M&R capability, railway managers can determine the track grids necessary for M&R activities and provide decision support for optimal scheduling of M&R.

Future work should further study and expand in two directions: (1) gradually improve and perfect the condition evaluation index system of track grids; (2) optimize the scheduling of M&R based on the track grid health index.

References

1. Railway, E.C.E.C. 2004. *Encyclopedia of china railway engineering & railway maintenance*. Beijing: China Railway Publishing House.
2. Gao, L. 2010. *Track engineering*. Beijing: China Railway Publishing House.
3. Vesković, S., J. Tepić, M. Ivić, G. Stojić, and S. Milinković. 2012. Model for predicting the frequency of broken rails. *Metalurgija* 51: 221–224.
4. Sudafed, D.H., and C.A. Barkan. 2008. Hybrid logistic regression/neural network model for the prediction of broken rails. In *Proceedings of the 8th World Congress on Railway Research, Seoul, Korea*.
5. Sadeghi, J.M., and H. Askarinejad. 2011. Development of track condition assessment model based on visual inspection. *Structure and Infrastructure Engineering* 7: 895–905.
6. Bai, L., R. Liu, Q. Sun, F. Wang, and F. Wang. 2016. Classification-learning-based framework for predicting railway track irregularities. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 230: 598–610.
7. Xin TFSM. 2015. Grey-system-theory-based model for the prediction of track geometry quality. In *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit*.
8. Bai, L., R. Liu, Q. Sun, F. Wang, and P. Xu. 2015. Markov-based model for the prediction of railway track irregularities. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 229: 150–159.
9. Kang, X., W. Wang, and J. Liu. 2013. Research on comprehensive evaluation system for track irregularity of high-speed railway based on RAMS. *China Railway Science* 34:13–17.

10. British-Standards-Institution. 1999. EN 50126—1999 railway applications: The specification and demonstration of reliability, availability, maintainability and safety. London: British Standards Institution.

11. Sun, L., and W. Gu. 2010. Pavement condition assessment using fuzzy logic theory and analytic hierarchy process. *Journal of Transportation Engineering* 137: 648–655.

12. Koduru, H.K., F. Xiao, S.N. Amirkhanian, and C.H. Juang. 2010. Using fuzzy logic and expert system approaches in evaluating flexible pavement distress: Case study. *Journal of Transportation Engineering* 136: 149–157.

13. Zhang, W., K. Sun, C. Lei, Y. Zhang, H. Li, and B.F. Spencer. 2014. Fuzzy analytic hierarchy process synthetic evaluation models for the health monitoring of shield tunnels. *Computer-Aided Civil and Infrastructure Engineering* 29: 676–688.

14. Bianchini, A. 2012. Fuzzy representation of pavement condition for efficient pavement management. *Computer-Aided Civil and Infrastructure Engineering* 27: 608–619.

15. Haas, R., W.R. Hudson, and J.P. Zaniewski. 1994. *Modern pavement management*. Malabar, FL: Krieger.

16. Liang, M.T., T.B. Chu, W.H. Tsao, and C.J. Yeh. 2006. Determining the repair ranking of existing RC bridges using fuzzy synthetic evaluation method. *Journal of the Chinese Institute of Engineers.* 29: 37–50.

17. Zayed, T., Z. Chen, and A. Qasem. 2014. Infrastructure performance rating models for wastewater treatment plants. *Structure and Infrastructure Engineering*. 1–15.

18. Liu, R., L. Bai, F. Wang, Q. Sun, and F. Wang. 2015. Grid: A new theory for high-speed railway infrastructure management. *Transportation Research Board 94th Annual Meeting*. Washington DC, United States.

19. China TRMO. 2010. Railway Transport [2006] No. 41. Rules of railway track maintenance.

20. EN 13848-1-2008 Railway Applications Track—Track Geometry Quality—Part 1 Characterisation Of Track Geometry.

21. China TRMO. 2009. Transport base line [2009] No. 41. Track quality index management interim standards of existing railway.

22. Jolliffe I. 2002. *Principal component analysis*. Hoboken, New Jersey: Wiley Online Library.

23. Hartigan JA和W. 1979. Algorithm AS 136: A K-Means Clustering Algorithm. *Journal of the Royal Statistical Society. Series C (Applied Statistics)* 28:100–108.

24. R Core Team. 2013. R: A language and environment for statistical computing. Available at: http://www.R-project.org/.

25. Lê, S., J. Josse, and F. Husson. 2008. FactoMineR: An R package for multivariate analysis. *Journal of Statistical Software* 25: 1–18.

Efficiency Evaluation Model of Car Sharing for Low-Income People

Liya Yao and Kai Chen

Abstract The newly emerged traffic mode, car sharing, is helpful to reduce energy consumption and traffic pressure. Low-income people, who have fewer options for transportation and little choice in terms of employment location, need to be paid more attention. In order to evaluate travel efficiency of car sharing system, this paper studied the commuting travel characters of low-income people, including their trip time, trip mode choice, trip length, trip frequency, and route choice each day. Then we put forward the evaluation indexes and studied the efficiency evaluation model of car sharing system and other traffic choices. Considering the characters of various traffic modes, we developed an efficiency index which including the total travel time, distance, and cost. Efficiency evaluation model of car sharing was built using the investigation data of travel mode selection. The results show that car sharing is an efficient traffic mode during peak hours in Beijing. The results are meaningful to guide the operation and management of urban public transport.

Keywords Efficiency evaluation · Low-Income People · Car sharing · Data Envelopment Analysis

Introduction

Along with economic growth, ownership of motor vehicles and traffic volume increased rapidly. This has put significant pressure on energy supply, traffic management, environmental protection, etc. Reducing the use of private vehicles has become an important factor in solving the above problems. In order to alleviate the negative impact of cars on the environment, some European and American countries have developed innovative automobile consumption patterns based on networking and other methods, such as car sharing. Car sharing is a new travel

L. Yao (✉) · K. Chen
Department of Transportation Engineering, Beijing Institute of Technology, 5 South Zhongguancun Street, Haidian District, Beijing 100081, People's Republic of China
e-mail: yaoliya@bit.edu.cn

method. People rent cars for short period of time, often by the hour. Customers who make only occasional use of a vehicle, as well as others who would like to try a different type than they use day to day are interest in car sharing. In order to compare the cost and benefit of the traffic modes (private car, subway, bus, taxi, bicycle, walking, and car sharing), an efficiency evaluation model is necessary.

Literature Review

In the past, scholars at home and abroad have achieved great progress in research on travel characteristics of low-income people.

First, wages are negatively associated with distance. Paul [1] established set of models to describe the role of job access and spatial mobility in improving the employment opportunities of welfare recipients. He found that people who live in job-rich neighborhoods are more likely to find jobs near their homes. Similar conclusion was drawn by Blumenberg [2]. She examined changes in the travel behavior patterns of the working poor based on descriptive data and found that rural, transit-dependent welfare recipients have only limited access to employment opportunities within a reasonable commute distance. Furthermore, Blumenberg [3] examined changes in the travel behavior patterns of the working poor based on descriptive data and found that low-income people become increasingly reliant on public transit for access to work. Besides this, Scholl (2002) reviewed the literature on travel patterns of low-income populations. She found that these patterns may be as reflective of the unique needs of low-income groups as they are of what is lacking in terms of transportation options in comparison with those available to higher income groups.

From all the above researches, it has been found that the travel mode of the poor is an important issue in the development of public transportation system in big cities. As the newly emerged traffic mode, car sharing, which is an efficient traffic mode for the poor, is becoming more and more popular.

Shaheen [4] summarized the developing history of car sharing in different countries. The benefits brought by car sharing were analyzed, including energy and cost savings, environmental protection, and so on. Advantages and experiences of car sharing in several countries were compared. And Chen [5] analyzed the effect of car sharing on travel behavior, infrastructure demand, and other traffic modes. The author examined the life-cycle impact of car sharing and compared the life-cycle reductions in energy and GHG emissions of car sharing and private vehicles.

In China, car sharing is a newly emerging traffic mode, which needs guidance in its development from both government and passengers. Because of the particularity of the urban traffic in China, whether car sharing is suitable for low-income people is a question worth exploring.

In order to present an efficiency evaluation model of car sharing, and provide a base for the management and operation for car sharing, this paper describes an efficiency evaluation model of car sharing based on the total efficiency of passengers.

This paper is structured as follows: Section "Efficiency Evaluation Model of Car Sharing in Big Cities of China" discusses a basic method of evaluating car sharing efficiency and surveys the input and output indexes. Section "Date Analysis" discusses the travel characters of all the people, especially low-income households based on trip data of the households in Beijing, followed by the conclusions in Section "Efficiency Evaluation Model of Car Sharing for Low-Income People."

Efficiency Evaluation Model of Car Sharing in Big Cities of China

Basic Method

Data envelopment analysis (DEA) is a methodology, developed by Charnes et al. [6], for measuring the relative efficiency of peer decision making units (DMUs) that have multiple inputs and outputs.

The basic method of DEA [6] is as formula (1).

$$E_k = \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{ik}} \tag{1}$$

Suppose that there are n kinds of decisions in total and decision making units (DMUs) denoted by U_k ($k = 1, 2, \ldots, n$) to be analyzed. Consumes amount x_{ik} of input i and produces y_{rk} of output r. x_{ik} denotes the wastage of decision making unit k to resource i, and y_{rk} denotes the sendout of decision making unit k to output r.

DEA optimization treats the observed vectors of inputs X_k and outputs Y_k as given and chooses values of input and output weights for particular u_r and v_i with respect to the following inputs minimizing optimization problem [6]:

$$\left. \begin{array}{c} \max \left[E_{k_0} = \frac{\sum_{r=1}^{s} u_r y_{k_0}}{\sum_{i=1}^{m} v_i x_{k_0}} \right] = V_P \\ \text{s.t.} \quad E_k = \frac{\sum_{r=1}^{s} u_r y_{rk}}{\sum_{i=1}^{m} v_i x_{rk}} \leq 1, \quad k = 1, 2, \ldots, n \\ u_r, v_i \geq 0, \forall r, i \end{array} \right\} \tag{2}$$

where u_r is the weight for output r and v_i is the weight for input i.

Comparing with AHP and FCE

Analytic hierarchy process (AHP) and the fuzzy comprehensive evaluation method (FCE) are both multi-rule evaluation methods used for the comprehensive systems whose data are difficult to obtain.

1. AHP method

AHP is a model for the study of the problem of nonstructural decision making strategy by using the combination of qualitative and quantitative indicators. The main core of this method is to establish a multilevel hierarchical structure model, and its essence is a kind of decision making method for comprehensive studies by weighing the target relationship.

AHP method has its unique characteristics and application in that it provides a system for the systems analyst a systematic, model-based way of thinking and system synthesis. The program takes target weights for the calculation as standards, to achieve qualitative and quantitative thinking. Therefore, this method can study on the implementation of the multilevel classification, semi-structured and unstructured problems.

(2) FCE method

Fuzzy comprehensive evaluation is a comprehensive evaluation method based on fuzzy mathematics. The comprehensive evaluation method based on the theory of fuzzy membership of the qualitative evaluation into quantitative evaluation, which uses fuzzy mathematics by a variety of factors things or objects to make an overall evaluation. It has a clear result, systematic and strong features. Therefore, it can solve vague problems which are difficult to quantify.

However, AHP and FCE are both limited to each fragmented industry sector which measure the efficiency under different conditions from the comparison evaluation department, rather than comparing the various sectors under the same conditions. Since this measure does not have comparable efficiency, greatly affected the model to evaluate the validity of the conclusions. DEA method views system metrics to determine the efficiency of various departments, which is its advantage.

Evaluation Index System of Car Sharing

It is necessary to evaluate the efficiency of car sharing before analyzing the actual effects of the new traffic mode, car sharing. Taking all the possible indexes of input and output, and considering the conveniences of quantifying, evaluation index system of car sharing efficiency is shown in Fig. 1.

Date Analysis

Definition of Low-Income People

This study focused on the travel characteristics of low-income inhabitants. According to the China Statistical Yearbook 2014, low-income people were defined
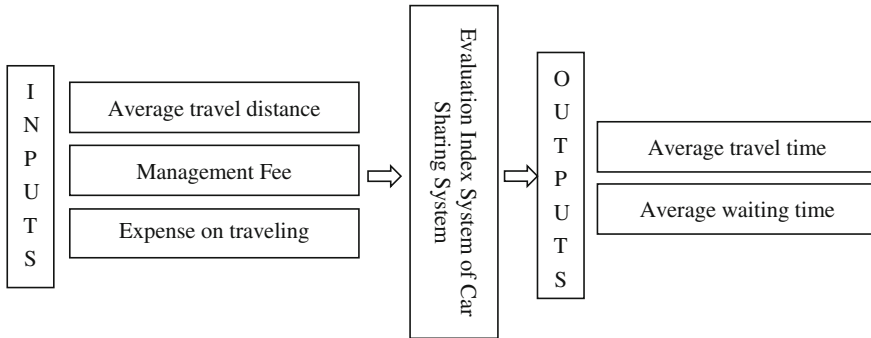
Fig. 1 Evaluation index system of car sharing system

as those 20% people who fell in the bottom by per capita disposable income quintile, as shown in Tables 1 and 2.

Taking this as the low-income standard, low-income people were selected. There are 777 samples in trip file and 454 are low-income samples. Low-income samples occupy about 58.43% of the total.

Table 1 Per capita disposable income of urban households by income quintile

Year	Low-income households	Lower middle income households	Middle income households	Upper middle income households	High-income households
2000	3132.0	4623.5	5897.9	7487.4	11,299.0
2001	3319.7	4946.6	6366.2	8164.2	12,662.6
2002	3032.1	4932.0	6656.8	8869.5	15,459.5
2003	3295.4	5377.3	7278.8	9763.4	17,471.8
2004	3642.2	6024.1	8166.5	11,050.9	20,101.6
2005	4017.3	6710.6	9190.1	12,603.4	22,902.3
2006	4567.1	7554.2	10,269.7	14,049.2	25,410.8
2007	5364.3	8900.5	12,042.2	16,385.8	29,478.9
2008	6074.9	10,195.6	13,984.2	19,254.1	34,667.8
2009	6725.2	11,243.6	15,399.9	21,018.0	37,433.9
2010	7605.2	12,702.1	17,224.0	23,188.9	41,158.0
2011	8788.9	14,498.3	19,544.9	26,420.0	47,021.0
2012	10,353.8	16,761.4	22,419.1	29,813.7	51,456.4
2013	11,433.7	18,482.7	24,518.3	32,415.1	56,389.5

Data from http://www.stats.gov.cn/tjsj/ndsj/2014/zk/html/Z0607C.html

Table 2 Per capita disposable income of urban households

Year	Low-income households	Lower middle income households	Middle income households	Upper middle income households	High-income households
2014	20,154.7	31,703.1	4478.2	59,717.6	87,380.9

Travel Characters of Low-Income People

Usually, traffic mode and traffic frequency are closely related to the total income of households. Households with high incomes generally travel by car or taxi. In contrast, households with low incomes tend to travel by bus or other modes. People with high income make more travels than people with low income.

Based on the data of all the low-income people, travel characters were found as follows.

(1) Distribution of traffic mode by income

Making data collation on the distribution of traffic income, as shown in Fig. 2. For low-income families, the bus is the highest proportion of traffic mode while car is the lowest proportion (Figs. 3, 4 and 5).
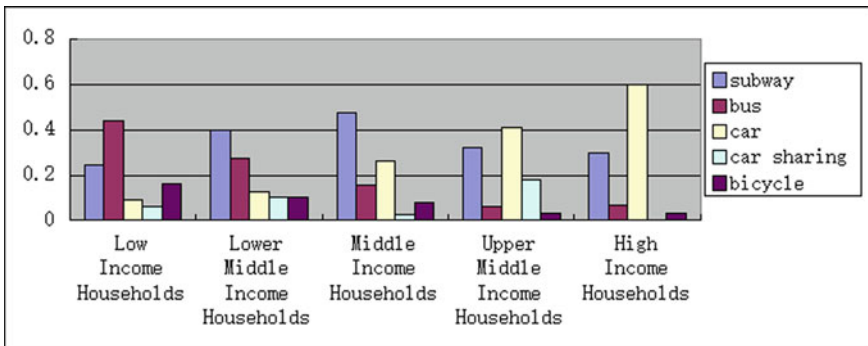
Fig. 2 Distribution of traffic mode by income
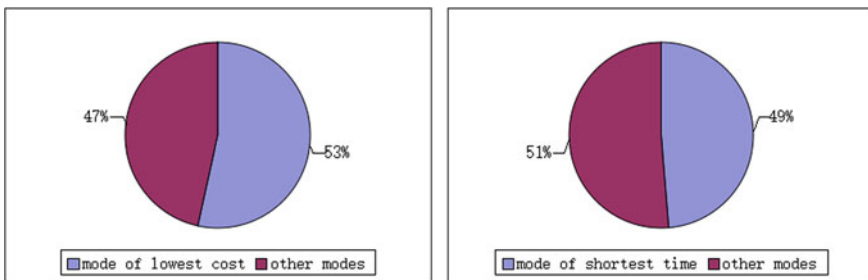
Fig. 3 Distribution of the lowest cost or the shortest time

			Mode choice	Lowest cost
Spearman's rho	Mode choice	Correlation Coefficient . .	1.000	.637
		Sig. (2-tailed)		.002
		N . .	454	454
	Shortest time	Correlation Coefficient . .	.637	1.000
		Sig. (2-tailed)	.002	
		N . .	454	454

Fig. 4 Spearman correlation analysis between mode choice and mode of lowest cost

			Mode choice	Lowest cost
Spearman's rho	Mode choice	Correlation Coefficient . .	1.000	.346
		Sig. (2-tailed)		.002
		N . .	454	454
	Shortest time	Correlation Coefficient . .	.346	1.000
		Sig. (2-tailed)	.002	
		N . .	454	454

Fig. 5 Spearman correlation analysis between mode choice and mode of shortest time

(2) Influence by time and fee

To study the influence factors on the distribution, data are further processed. According to the correlation analysis, it is found that the final choice of travel mode is mostly related to expenses and time when modes are optional. For low-income people, weighs of expenses should be considered higher.

Efficiency Evaluation Model of Car Sharing for Low-Income People

Evaluation Index System

The efficiency of car sharing reflects the virtues or defect degree on average travel time, average waiting time under constraint of average travel distance, expense on traveling, and management fee. As shown in Fig. 1, inputs and outputs are as follows.

Inputs: average travel distance, expense on traveling, and management fee.
Outputs: average travel time and average waiting time.

Calculation of Evaluation Indexes

All parameters required are derived from the data processing of the survey sample.

(1) Average travel distance: Average travel distance refers to the sample mean of the traveling distance for low-income people. It can be represented by mathematical expression below, in which d_i represents the travel distance of sample i.

$$\frac{1}{N} \sum d_i \tag{4}$$

(2) Expense on traveling: Expense on traveling refers to the average cost of single trip for different travel mode selection. The contents of the cost can be different. For instance, the cost of traveling by car contains the cost of gasoline and parking fee, while the cost of traveling by bicycle can be zero.
(3) Management Fee: Management Fee refers to the average value the operational costs in per unit of time, including purchase and maintenance costs of vehicles, stuff salaries.
(4) Average travel time: Average travel distance refers to the sample mean of the single-trip time for low-income people. Single-trip time is the time from the start to the end of the ride, waiting time on the way, such as waiting for the traffic lights should be included.
(5) Average waiting time: Average travel distance refers to the sample mean of the waiting time before customers are picked up by the vehicles.

Calculation of Efficiency Evaluation Model

Based on Table 3, we got the efficiency score and ordering results of the different modes of traffic selections, as shown in Table 4.

Results

For the normal seven travel methods, Table 4 summarizes the efficiency scores such as the average travel distance, average travel time, and travel expense. According to the calculation results, car sharing is the top three travel options. However, the highest score does not necessarily indicate that the indicator is the most important. The system of car sharing can obviously be optimized by reducing the customer waiting time.

Table 3 Raw data of the different modes of traffic selection

	Average travel distance {I}	Expense on traveling {I}	Management fee {I}	Average travel time {O}	Average waiting time {O}
Private car	7.3	12.95	128.63	16.85	1
Subway	10.9	6.53	0	33.48	4.37
Bus	12.77	2.15	0	62.2	24.7
Taxi	6.33	27.44	0	25	5.37
Bicycle	3.98	0	17.57	37.34	1
Walking	1.91	0	0	26	1
Car sharing	8.21	10.35	0	16.56	3.89

Table 4 Raw data of the different modes of traffic selection

DMU	Score (%)	Average travel distance {I} {V}	Expense on traveling {I}{V}	Management fee {I}{V}	Average travel time {O}{V}	Average waiting time {O}{V}
Private car	85.83	0.18	0.12	0.15	0.32	0.09
Subway	96.74	0.25	0.22	0.32	0.13	0.05
Bus	91.44	0.27	0.26	0.32	0.05	0.01
Taxi	84.33	0.10	0.12	0.32	0.26	0.03
Bicycle	76.41	0.08	0.33	0.23	0.03	0.09
Walking	79.27	0.03	0.33	0.32	0.02	0.09
Car sharing	96.44	0.22	0.15	0.32	0.21	0.06

Conclusions

Evaluation model of car sharing is established based on the investigation data of 454 trips of low-income people during peak hours in Beijing. Operational efficiency of different traffic modes can be evaluated by the overall performance indicator. Furthermore, the potential improvements of the efficiency indicators can be calculated, which is helpful for optimizing urban public traffic. Thus the model developed in this paper can be used to forecast the travel demand of each mode and to form a base for transportation facility design and management.

References

1. Ong, Paul, and Evelyn Blumenberg. 1998. Job access, commute and travel burden among welfare recipients. *Urban Studies* 35(1): 77–93.
2. Blumenberg, Evelyn, and Kimiko Shiki. 2003. How welfare recipients travel on public transit, and their accessibility to employment outside large urban centers. *Transportation Quarterly* 57(2): 25–37.
3. Blumenberg, Evelyn, and Trevor Thomas. 2013. Travel behavior of the poor post-welfare reform. In *TRB 2014 Annual Meeting.*
4. Shaheen, S., D. Sperling, and C. Wagner. 1998. Carsharing in Europe and North America: Past present and future. *Transportation Quarterly* 52(3): 35–52.
5. Chen, T. Donna, and Kara M. Kockelman. 2015. Carsharing's life-cycle impacts on energy use and greenhouse gas emissions. *Transportation Research Record*, January 2015.
6. Charnes, A., W.W. Cooper, and E. Rhodes. 1978. Measuring efficiency of decision-making units. *European Journal of Operational Research* 2: 429–444.
7. Bardhi, Fleura, and Giana M. Eckhardet. 2012. Access-based consumption: The case of car sharing. *Journal of Consumer Research* 39.
8. Barnum, D.T., J.M. Gleason, and B. Hemily. 2008. Using panel data analysis to estimate DEA confidence intervals adjusted for the environment. *Journal of Transportation Engineering, ASCE* 134 (5): 215–223.
9. Huijberts, H.J.C. 2002. Synchronous motion in a continuous car-following model for a bus route. *Proceedings of the IEEE Conference on Decision and Control* 1: 716–721.
10. Min, Hokey, and Byung In Oark. 2005. Evaluating the inter-temporal efficiency trends of international container terminals using data envelopment analysis. *International Journal of Integrated Supply Management, International Journal Supply Management (Switzerland)* 1(3), 258–277.

Railway Energy Consumption Analysis Based on Regression Model

Liang Sun, Ling Lin, Liang Chen, Mei Liu, Wei Bao and Li Wang

Abstract The energy consumption and related influence factors of railway transportation are analyzed based on the statistical data of national railway. Then, two transportation energy consumption evaluation indicators, the comprehensive energy consumption per unit service and the emissions of sulfur dioxide, are conducted the correlation analysis, which can derive the estimated emissions of sulfur dioxide between 2003 and 2006 that has been missing. Based on the study, two multi-variable linear regression models are formulated: One is comprehensive energy consumption of transportation, and another is the emission of sulfur dioxide. After the analyzation, the result shows that the two important indicators of energy consumption show significant linear relationship with both passenger turnover and freight turnover. According to this, the work predicts the energy consumption of China National Railway in 2016.

Keywords Railway transportation · Energy consumption · Factor identify · Regression analysis

The world pays great attention to transportation energy saving in a variety of methods and measures. At present, the infrastructure construction of railway transportation has been developing at a high speed in China; the operation mileage of high-speed railway is more than 60% of the world's total mileage. Under the background of a wide range of network operation, the energy consumption of railway transportation has been paid more attention [1]. Shi et al. analyzed the influence of energy consumption from 7 elements of time, station, line, area, vehicles, equipment and season, and established the evaluation system of energy consumption [2]. Fan YongTao and Li Xia analyzed the impact of energy

L. Sun · L. Lin · L. Chen · M. Liu · W. Bao
China National Institute of Standardization, Beijing 100191, China
e-mail: sunliang@cnis.gov.cn

L. Wang (✉)
School of Traffic and Transportation, Beijing Jiaotong University, Beijing 100044, China
e-mail: wangli@bjtu.edu.cn

consumption on driving conditions, technical equipment, and three aspects of management level, and put forward some energy saving measures [3]. Wang Yuming analyzed the energy consumption influence and measures of line design, technology such as driving speed selection, marshaling technology, and strategy and application of new technology [4]. Sun et al. established the gray correlation degree model in order to quantitatively determine the factors of energy consumption intensity based on the identification of the organization factors of railway passenger transport energy consumption [5]. Based on the systematic analysis of energy consumption influence factors of railway transportation, this chapter analyzes the development trend of energy consumption from the two aspects of the national railway network infrastructure construction and the main transport task, and estimates the energy consumption of national railway transport through the regression analysis and correlation analysis.

Analysis of the Energy Consumption Factors of Railway Transportation

With the rapidly development of national high-speed railway in China, the scale of railway station enlarges constantly. Although the railway freight turnover declines continuously during last three years, the demand of the railway passengers increases substantially. The energy consumption of railway is only next to water transportation ranked fourth in the five kinds of transportations [3]. The main factor of the energy consumption of railway is to accomplish the various types of train operations for the demand of passengers and freights. Meanwhile, factors, such as the level of the transportation organization management and the level of the transportation equipment technic, can affect the railway energy consumption.

1. Effect of the train operation

   The energy consumption in the process of the train operation is due to various resistances of the train, the circuit conditions, the braking, the acceleration system of the train, and other kinds of train operation mainly. There are two kinds of resistances of the train operation: the mechanical resistance and the aerodynamic resistance. The mechanical resistance includes the bearing drag resistance, the wheel rolling resistance, the wheel slip resistance, and the shock and vibration resistance. The aerodynamic resistance mainly caused by the relative movement of the ambient air due to the train operation, which is relative to the speed, the shape, and the size of the train. It is considering that the basic resistance of the train operation is directly proportional to the square of the train speed generally, which means that the higher the speed is, the greater the driving resistance and the higher the energy consumption will be. Moreover, the circuit condition is an important factor for energy consumption. The large line slope and huge line transformation can not only cause the energy consumption to overcome the gravity of the trains,

but also increase the shock and vibration resistances. The emergencies often cause the waste of energy, such as the temporary parking and limited speed due to the emergency equipment failures or the boundary intrusion. To keep the safety of the train operation, the drivers need to operate the acceleration and the braking system frequently, which can cause more waste of energy.

2. Effect of the management level of the transportation organization

In the condition of the fixed transportation capacity, if the number of transport production was higher, then the energy consumption of per unit production is lower, which means that the efficiency of the transportation organization management is higher. However, the transportation organization management contains the transportation organization models, the operation scheme, the operation plan, the maintenance plan, the crew plan, and the contents of each level. For the freight transportation, it involves the loading plan, the freight plan, the shunting plan, and other contents. The transportation organization scheme with high efficiency can take full advantage of the network transportation capacity. In the same condition, improving the passenger turnover and the freight turnover can reduce the energy consumption of per unit transportation capacity efficiently.

3. Effect of the technical level of the transport equipment

The technology level of China's railway transportation equipment has improved a lot compared to the past. The electrified railway has taken over 60% of the national railway. The energy consumption can be affected by the efficiency of the key technical equipment, such as all kinds of basic signal devices, the operation blocking system, the station chain system, the traction control system, and the train control system.

Relative Analysis of the Energy Consumption of the Railway Transportation

Due to the absence of the observed data of the sulfur dioxide emissions of the national railway transportation, the relative analysis of the comprehensive energy consumption per unit transportation workload and the sulfur dioxide emissions at first are carried out. Then, the two energy consumption indexes are estimated by using the regression analysis.

Based on the national railway statistical bulletin of the year 2003–2015, by the conversion of data in the bulletin, the passenger capacity, the passenger turnover, the freight capacity, the freight turnover, the operating mileage, the electrification mileage, the comprehensive energy consumption per unit transportation workload, the data of the sulfur dioxide emissions are obtained as shown in Table 1.

Correlation analysis is to study whether there is some kind of interdependent relationship between the research objects, and to analyze the relative level between

Table 1 Energy consumption and related factors of railway

Year	Passenger capacity (10^2 million person)	Passenger turnover (Trillion person km)	Freight capacity (10^2 million tons)	Freight turnover (Trillion tons)	Operation mileage (10,000 km)	Electrification mileage (10,000 km)	Comprehensive energy consumption per unit transportation workload (tce/millions of conversion ton-km)	Sulfur dioxide emissions (10,000 tons)
2015	25.35	1.20	33.58	2.37	12.10	7.40	4.68	2.88
2014	23.05	1.12	38.13	2.75	11.20	6.50	4.55	3.17
2013	21.06	1.06	39.66	2.92	10.31	5.58	4.66	3.53
2012	18.93	0.98	39.04	2.92	9.76	5.10	4.74	3.78
2011	18.62	0.96	39.32	2.95	9.33	4.61	4.76	4.01
2010	16.76	0.87	36.42	2.76	9.12	4.25	5.01	4.03
2009	15.25	0.79	33.20	2.49	8.60	3.60	5.30	4.02
2008	14.61	0.78	32.59	2.48	8.00	2.80	5.60	4.23
2007	13.17	0.72	31.45	2.40	7.80	2.55	5.78	4.36
2006	12.57	0.66	28.71	2.17	7.70	2.44	6.12	–
2005	11.56	0.61	26.83	2.05	7.54	2.02	6.48	–
2004	11.18	0.57	24.81	1.91	7.44	1.93	6.65	–
2003	9.73	0.48	22.35	1.71	7.30	1.88	7.32	–

the variables through the certain technical index. This chapter adopts Pearson's sample correlation coefficient (formula 1) to analyze the correlation between the comprehensive energy consumption per unit transportation workload and the sulfur dioxide emissions, and uses the t statistic to process the significant test, as shown in formula (2).

$$r = \sum_{i=1}^{n} (x_i - \bar{x})(y_i - \bar{y}) \left/ \sqrt{\sum_{i=1}^{n} (x_i - \bar{x})^2 \sum_{i=1}^{n} (y_i - \bar{y})^2} \right. \tag{1}$$

$$t = r\sqrt{n-2} \left/ \sqrt{1 - r^2} \right. \tag{2}$$

Setting up significant level as 0.01 in the statistical test, according to the calculation of the data in Table 1, it is noticeable to obtain Pearson's correlation coefficient, which is 0.841, between the comprehensive energy consumption per unit transportation workload and sulfur dioxide emissions. The associated probability of statistical tests, which is 0.002, is much smaller than significant level; therefore, it is considered that the relationship between the comprehensive energy consumption per unit transportation workload and sulfur dioxide emissions shows a strong positive correlation relationship.

On the basic of the determined relevant relations, the relationship between the comprehensive energy consumption per unit transportation workload and the sulfur dioxide emissions conformed to formula (3) by processing the curve fitting. The sulfur dioxide emissions are taken as independent variables, and the comprehensive energy consumption per unit transportation workload is taken as the dependent variable, then the goodness of fit is 0.919. The curve and data scatter diagram are shown in Fig. 1.

$$y = 13.958 - 5.798 \cdot x + 0.893 \cdot x^2 \tag{3}$$

According to formula (3), the sulfur dioxide emissions from the year 2003 to 2006 could be estimated as 4.38, 4.54, 4.61, and 4.85 separately, with the unit 10,000 tons.

Regression Model of the Energy Consumption of the Railway Transportation

In this section, it adopts regression analysis method to set up the regression model of transport energy consumption from the China National Railway infrastructure construction and transportation capacity.

Regression analysis is mainly used to study the correlation between variables. First of all, assume that the relationship between the dependent variables and
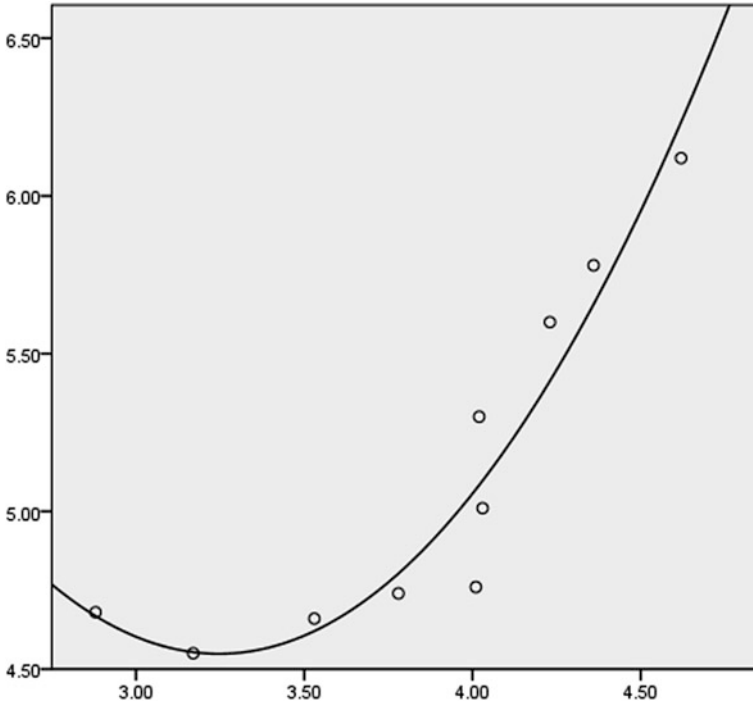
Fig. 1 Fitting curve and scatter chart of data points

independent variables is linear. Then, a linear regression model is used to fit the dependent variable data and independent variable data. In addition, the regression equation is obtained by determining the model parameters. After that, we could analyze the correlation relationship between the variables through the regression equation. In this section, it adopts the multivariate linear regression to study the correlation between the passenger capacity, passenger turnover, freight capacity, freight turnover, mileage, electrified mileage, comprehensive energy consumption per unit transportation workload, and the sulfur dioxide emissions, assuming that both of the regression equations are satisfied in the form of formula (4).

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_i x_i + \cdots + \beta_k x_k + \varepsilon \tag{4}$$

After determining the parameters, the regression model needs to be proceeded with the significant test. Firstly, we need to process the F test to the significance of multivariate linear regression equation. Testing procedure shows below: taking the original hypothesis as H_0: $\beta_i = 0$, if H_0 was accepted, it suggests that the relationship between the dependent variable y and the independent variable x_i is inappropriate to described by using the linear regression model. Using the decomposition of the total sum of the square of dispersion constructs the F test, as shown in formula (5).

$$F = \sum (\hat{y} - \bar{y})^2 / k \Big/ \sum (\bar{y} - \hat{y})^2 / (n - k - 1) \tag{5}$$

where $\bar{y}$ stands for the sample mean, $\hat{y}$ stands for the predicted value of the samples, n stands for sample number, k stands for the number of the independent variables, the statistical magnitude of F obeys the distribution first degree of freedom of k, and the second degree of freedom is $n - k - 1$ for F distribution. If the value of F is greater, it declares that the effect due to changes in the dependent variables caused by the independent variables is greater than the random factors. Therefore, the significant level α can be given. The corresponding concomitant probability of F can be calculated as the value p. If $p < \alpha$, then the original assumption H_0 is rejected; considering that it is linear relationship between the dependent variables and the independent variables, the regression equation is significant and otherwise, the regression equation is not significant. In addition, it is required to take the significant test of the regression coefficients, which means to test the influence of each independent variable on the dependent variables in order to find out factor x, which affects y significantly. It generally uses the t test for the significant test of regression coefficients. The process is similar to the inspection of the regression equation above.

Regression Analysis of the Comprehensive Energy Consumption Per Unit Transportation Workload

First of all, the passenger capacity, passenger turnover, freight capacity, freight turnover, operating mileage, and electrification mileage are assumed as the independent variables, signed as x_1, x_2, x_3, x_4, and x_5 separately, using the least square method to calculate the regression equation of the comprehensive energy consumption of unit transportation workload, as shown in formula (6).

$$y_1 = 8.915 + 0.227x_1 - 7.429x_2 + 0.034x_3 - 0.985x_4 + 0.119x_5 - 0.176x_6 \tag{6}$$

According to the method above, the regression equation significant test to the formula (6) is proceeded, obtaining the value $F = 65.2$ and its associated probability $p < 0.001$, which suggests that it is significant linear relationship between the comprehensive energy consumption per unit transportation workload and the factors of the passenger capacity, passenger turnover, freight capacity, freight turnover, operating mileage, and electrification mileage. The regression equation is significant.

Second, t test is used to analyze the regression coefficient, as in Table 2.

From the data of the analysis, Table 2, it shows that the significant probabilities of each coefficients in the formula (6) are all greater than 0.001. Therefore, in this chapter, it removed the relative factors successively, according to the significant value from large to small, and recalculates the regression model. At the meantime,

Table 2 Table for regression coefficient and significant test of comprehensive energy consumption

Coefficients	T	Significance
(Constant)	1.669	0.146
Passenger capacity (10^2 million person)	0.732	0.492
Passenger turnover (trillion person km)	−1.437	0.201
Freight capacity (10^2 million tons)	0.114	0.913
Freight turnover (trillion tons)	−0.267	0.799
Operation mileage (10,000 km)	0.135	0.897
Electrification mileage (10,000 km)	−0.271	0.795

the significant test is proceeded. Finally, it can obtain the formula (7), which is the regression equation of the comprehensive energy consumption per unit transportation workload. Here, the value of F changes from 65.200 to 271.335 with all three significant probabilities smaller than 0.001, which means that there is a strong correlation between the comprehensive energy consumption per unit transportation workload and the freight turnover as well as the passenger turnover.

$$y_1 = 10.022 - 2.326x_2 - 1.051x_4 \tag{7}$$

Regression Analysis of Sulfur Dioxide Emissions

Repeating the above process, first of all, the passenger capacity, passenger turnover, freight capacity, freight turnover, operating mileage, and electrification mileage are assumed as the independent variables, using the least square method to calculate the regression equation of the sulfur dioxide emissions, as shown in formula (8).

$$y_2 = 8.004 + 0.193x_1 - 4.890x_2 - 0.231x_3 + 3.203x_4 - 0.486x_5 + 0.252x_6 \tag{8}$$

According to the method mentioned above, the regression equation significant test to the formula (6) is proceeded, obtaining the value $F = 87.619$ and its associated probability $p < 0.001$, but proceed the t test to the regression coefficients obtains the Table 3. It shows that the significant probabilities of each coefficients in formula (8) are all greater than 0.001. Therefore, in this paper, it removed the relative factors successively, according to the significant value from large to small, and recalculates the regression model. At the meantime, the significant test is proceeded. Finally, it can obtain formula (9), which is the regression equation of the comprehensive energy consumption per unit transportation workload. Here, the value of F changes to 280.427, which is a huge increase, with all three significant probabilities smaller than 0.001. It suggests that there is a strong correlation between the sulfur dioxide emissions and the freight turnover as well as the passenger turnover.

Table 3 Table for regression coefficient and significant test of sulfur dioxide

Coefficients	T	Significance
(Constant)	2.736	0.034
Passenger capacity (10^2 million person)	1.139	0.298
Passenger turnover (trillion person km)	−1.727	0.135
Freight capacity (10^2 million tons)	−1.428	0.203
Freight turnover (trillion tons)	1.585	0.164
Operation mileage (10,000 km)	−1.002	0.355
Electrification mileage (10,000 km)	0.709	0.505

$$y_1 = 5.501 - 3.134x_2 + 0.461x_4 \qquad (9)$$

According to the railway construction planning, freight business adjustment, and analysis of historical data, passenger turnover and the freight turnover are expected to be 1.29 trillion passenger-kilometer and 2.05 trillion tons separately in 2016. According to formula (7) and formula (9), the national railway comprehensive energy consumption per unit transportation workload and sulfur dioxide emissions are predicted to 4.86 tce/millions of conversion ton-km and 24,000 tons separately in the year 2016.

Conclusion

This paper analyzes the related factors affecting the energy consumption of railway transportation and points out the train operation process, transportation organization, and management level, and the key technology of railway transportation equipment is the major influence factors to the energy consumption of the railway transportation. Moreover, according to the data of the calendar year national railway bulletin, this chapter analyzes the two key assessment indexes: the comprehensive energy consumption per unit transportation workload and sulfur dioxide emissions of the energy consumption of the railway transportation. Meanwhile, it estimated that the sulfur dioxide emissions from the year 2003 to 2006 were missing.

On the basis of this study, it establishes two multivariable linear regression models of the comprehensive energy consumption per unit transportation workload and sulfur dioxide emissions. The analysis shows that the two important indexes of energy consumption are both in the linear relationship with the passenger turnover and the freight turnover. As the turnover contains the two factors of capacity and the turnover mileage, the regression models conform to the real process of the railway transportation. Besides, this paper predicts the energy consumption of China National Railway Transportation in 2016 by using those regression models.

References

1. Jia, Shunping., Hongqin, Peng., and Shuang, Liu., et al. 2009. Review of Transportation and Energy Consumption Related Research. *Transportation System Engineering and Information* 9 (3): 11–15.
2. Shi, JIngya., Yongqing, Su., and Jiguang, Yue. 2008. Factors Analysis of Energy Consumption of Railway Transportation and Establish of Energy Consumption Estimation System. *Railway Transport and Economy* 30 (9): 46–49.
3. Fan, Yongtao, and Xiamiao Li. 2009. Analysis of Energy Consumption of Railway Transportation and Energy Saving Measures Research. *Management* 17 (6): 232–233.
4. Wang, Yuming, Mingjun Liu, and Yugang Zhao. 2011. Study on Energy-Saving Measures of Urban Rail Transit System. *Comprehensive Transportation* 4: 57–59.
5. Sun, Qipeng, and Cheng Xu. 2012. Analysis on Running Organization Factors and Gray Relation Grade Measure of Railway Energy Consumption. *Railway Transport and Economy* 34 (8): 56–61.

Research on Parking Choice Model Based on Shared Private Parking Space

Tongqiang Ding, Bo Wang, Lili Zheng, Jianfeng Xi, Shengli Wang and Shuangshan Xu

Abstract As the number of vehicles increases, the problem of parking has become more and more highlighted. This paper proposes the practice of sharing private parking space when not in use, to meet the ever-increasing parking demand through maximizing utilization of private parking lots. Optimal route and optimal parking location are two main focuses in this study on private parking sharing selection. Optimal route is a decision-making process based on the shortest travel time, using improved ant colony optimization algorithms to determine the route and travel time from travel origin to a shared private parking space, and to prepare for the quantification of the evaluation index for travel time in the shared parking space choice model, which first determines the quantified evaluation indexes according to the factors of interest when a person chooses a private parking space. Optimal parking location is determined by factoring in all the evaluation indexes and using the weighted TOPSIS model.

Keywords Shared parking · Parking choice · Route guidance · Ant colony optimization algorithms · TOPSIS

T. Ding · B. Wang · L. Zheng (✉) · J. Xi · S. Wang · S. Xu
College of Traffic, Jilin University, Changchun 130022, Jilin, China
e-mail: zlldtq1024@163.com

T. Ding
e-mail: dtq8@163.com

B. Wang
e-mail: 1357559062@qq.com

J. Xi
e-mail: xijf@jlu.edu.cn

S. Wang
e-mail: wsl20101130@163.com

S. Xu
e-mail: 1349519386@qq.com

Introduction

Private parking spaces are often at a lower utilization rate especially during daytime peak parking hours, all while public parking spaces, on the other hand, are fully occupied, which is a serious waste of parking resources [1]. With the advent of shared economy, idle resources are increasingly being shared among people. The idea for this paper also stems from the same ideology, the outcome of which is the idea of shared private parking [2]. Private parking space owners would also like to share their spare parking resources, but due to certain reasons, such as information asymmetry, limitation on foreign vehicles in private residential areas and so forth all, they prevent shared parking from becoming a reality [3].

February 21, 2016, the State Council of China issued the "CPC Central Committee and State Council on further strengthening urban planning and construction management," which decrees the gradual opening of enclosed residential areas and foreign vehicles having free access to parking resources in formerly restricted areas, making private parking space sharing a possibility.

Choice Factors for Shared Private Parking

Based on choice factors in private parking sharing, this chapter surveys some drivers in Changchun City of China and finds that people are most concerned about the following: parking safety, walking distance, parking convenience, parking cost, and parking accessibility, as shown in Fig. 1.

1. **Parking safety**

Parking safety refers to the safety a vehicle obtains after parking [4]. Vehicles parked on the roadside or at a non-secure location are subject to scratches, theft, and other incidents. In private parking spaces, safety can be worse than public parking
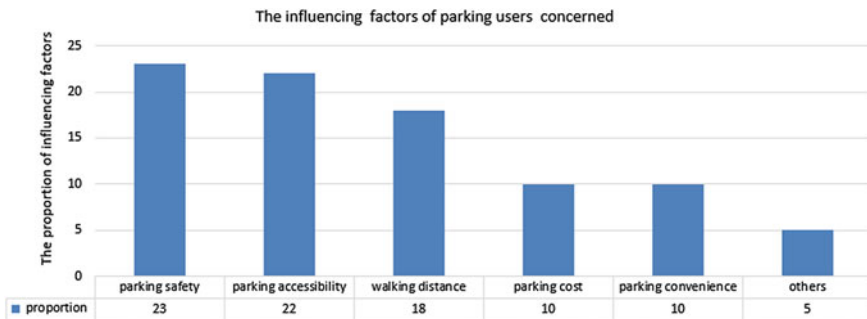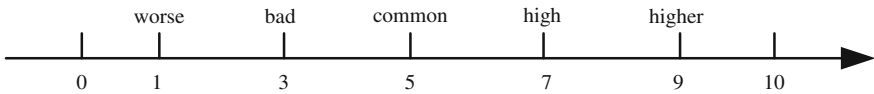

Fig. 1 Survey on parking decision-making factors

Fig. 2 Qualitative indicators of quantization

Table 1 Different types of parking security values

Parking type	Open-air lots	Underground lots	Mechanical lots	Residential underground lots
Security values	3–5	7–9	7–9	5–7

facilities, because there are no added protection measures that are common in public parking facilities; therefore, parking safety is the primary concern for parking users.

Parking safety and parking convenience are fuzzy evaluation factors and cannot be quantitatively analyzed,compared with walking distance, parking accessibility, and parking cost. To solve this problem, this paper maps the qualitative description to the coordinate axis, the evaluation of two qualitative indicators, as shown in Fig. 2.

According to the fuzzy efficiency evaluation of quantitative criteria, a questionnaire survey was conducted on different types of parking spaces, including open-air lots, underground lots, mechanical lots, and residential underground lots. The statistical analysis finds that safety for open-air parking space is the worst, safety for underground and mechanical lots is better, and safety for residential underground lots is moderate. Specific safety values are shown in Table 1.

2. **Parking accessibility**

Shared private parking accessibility refers to the difficulty for parking users from travel origin to shared parking space [5]. High parking accessibility suggests that the driver can easily reach the parking space, and it is a prerequisite for complete parking. Parking accessibility has a variety of selection criteria; for most parking users, they are usually most concerned about how to reach the destination in the shortest travel time; the most considered criterion is travel time.

In this paper, the calculation of travel time uses BPR impedance function by the USA, and it is calculated as follows:

$$t_i = \sum_{m=1}^{M} t_{im0} \left[1 + \alpha \left(\frac{Q_{im}}{C_{im}} \right)^{\beta} \right] \tag{1}$$

where Q_{im} represents the vehicle traffic flow of the m section to the ith private lot. C_{im} is the vehicle traffic capacity of the m section to the ith private lot. t_{im0} represents travel time of m section to the ith private lot when traffic volume is zero. α and β indicate regression parameters.

Suppose parking users reach the ith private lot after a total of M sections. For different locations of private parking space, the routes are not the same, and the route of number of M values may also be different.

3. Walking distance

Walking distance is the distance a parking user walks from the private lot to the destination. Walking distance is a factor parking users are more concerned with, which suggests the reachable convenience from the shared private lot to the destination [6]. February 21, 2016, the State Council issued the "CPC Central Committee and State Council on further strengthening urban planning and construction management," which will shorten the walking distance. Therefore, in order to simplify the calculation model for walking distance, we can use the formula for the straight-line distance between two points, and it is calculated as follows:

$$L_{ik} = \sqrt{(x_i - x_k)^2 + (y_i - y_k)^2} \tag{2}$$

where L_{ik} represents the walking distance from the ith private shared lot to the kth travel destination. (x_i, y_i) represents the ith private shared lot's position. (x_k, y_k) represents the kth travel destination's position.

4. Parking convenience

For shared parking convenience, parking type and parking forms need to be considered [7]. Different types of parking spaces have different parking forms. The three most common forms include diagonal parking, parallel parking, and vertical parking (as shown in Fig. 3).

Diagonal parking is the easiest to park, parallel parking is less easy, and vertical parking is the hardest. Different forms of parking values are shown in Table 2.

Fig. 3 Three common parking forms

Parallel parking

Vertical parkng

Diagonal parking
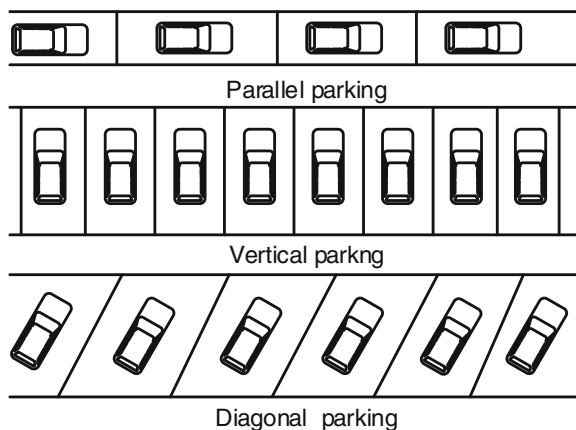
Table 2 Different forms of parking values

Parking forms	Slash form	Line form	Vertical form
Convenience values	7–9	4–6	1–3

5. **Parking cost**

Parking cost is the fee parking users need to pay to the private parking lot owner in order to park at his or her parking space. Parking cost affects directly the users' choice, and it is an important indicator on how users select parking location [8].

This paper divides cost into two charging methods: charges on shared-time and out-sharing period. Within the shared-time period, the price charges are fixed; out-sharing period requires charging high parking fees to compensate for the inconvenience to the private owners. It is calculated as follows:

$$C_{Pi} = \begin{cases} y_{i1} \cdot t_{Pi1}, & t_{Pi1} \in T_i \\ y_{i1} \cdot t_{Pi1} + y_{i2} \cdot t_{Pi2}, & t_{Pi2} \notin T_i \end{cases} \tag{3}$$

where C_{Pi} represents the ith private shared lot's parking charge. t_{Pi1} represents parking time in the shared-time period. t_{Pi2} represents parking time during out-sharing period. T_i represents shared time. y_{i1} represents the ith private shared lot's parking charge unit price in the shared-time period. y_{i2} represents the ith private shared lot's parking charge unit price in the out-sharing period.

Parking Choice Model for Shared Private Parking

Initial Selection

Basic private parking options include two stages: initial selection and final decision. The initial selection of private shared lot is based on two conditions, which are the maximum walking distance and sharing period.

Maximum walking distance is the distance parking users can withstand walking from the private parking space to the destination. Treating the destination as the center, the maximum walking distance as the radius, it is possible to provide service for parking users only in this range of the parking lot [9]. At the same time, the length of shared period needs to be taken into account to meet the parking user's need.

Alternative Optimal Route Selection

After the initial selection, there can be some alternative choices. In the next step, we need to calculate the optimal route and travel time from travel origin to each alternative private parking location [10]. After choosing travel route, travel time

also can be decided. This paper selects ant colony algorithm to solve the optimal routing problem. Ant colony algorithm, also known ant algorithm, is a method from imitating the behavior of ants searching for optimal routes [11]. Through passing information to each other, individual ants can obtain the optimal route from the ant nearest to the food source. The following introduces the state transition rule and pheromone update rule of ant colony algorithm [12].

1. **State transition rule**

$$P_{ij}^k(t) = \begin{cases} \dfrac{\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(t)}{\sum_{s\in \text{allowed}_k} \tau_{is}^{\alpha}(t)\eta_{is}^{\beta}(t)}, & j \in \text{allowed}_k \\ 0, & \text{others} \end{cases} \tag{4}$$

where P_{ij} presents choice probabilities ants move from node i to node j. allowed$_k$ represents the collection of allowed nodes for access. τ_{ij} represents the total amount of residual information between node i and node j. α represents the relative importance of the remnants. $\eta_{ij}(t)$ represents inspired factor, and β represents the relative importance of the expectation.

2. **Pheromone update rule**

Residual pheromone concentration changes with time, and τ_{ij} update formula is:

$$\tau_{ij}(t+n) = (1-\rho) \cdot \tau_{ij}(n) + \Delta\tau_{ij} \tag{5}$$

$$\Delta\tau_{ij}(t) = \sum_{k=1}^{m} \Delta\tau_{ij}^k(t) \tag{6}$$

$\Delta\tau_{ij}$ represents the toll of ant pheromone increment on the route from node i to node j. $\Delta\tau_{ij}^k(t)$ represents pheromone increment on the route from node i to node j of ant k.

$$\Delta\tau_{ij}^k(t) = \begin{cases} \mathcal{Q}/\mathcal{T}_k, & \text{ant } k \text{ pass } (i,j) \text{ in cycle } t \\ 0, & \text{others} \end{cases} \tag{7}$$

$\mathcal{Q}$ is constant, $\mathcal{T}_k$ represents the time length, ant k passes the route from travel start to end [13].

3. **Improved ant colony algorithm**

Ant colony algorithm is non-directional when the ant moves to the next node, which does not reflect the relationship between the next node and the end of the trip. This paper introduces the concept of gravitational pigment; gravitational pigment represents the attraction the travel end to the ants in the search route

process, the closer to the travel end, the greater attraction ants get, and vice versa. It can make the search of ants having the function of direction guidance, as shown in Fig. 4.

The gravitational pigment is related to the distance between the node and the target point. The gravitational influence of each node is calculated as follows:

$$F_j = \frac{G}{d_j} \tag{8}$$

F_j represents the gravitational pigment size from node j to the end of the trip. G represents the total amount of gravitational pigment for the trip endpoint. d_j represents the straight line from the node j to the end of the trip.

After increasing the concept of gravitational pigment, the inspired factor is modified as follows:

$$\eta_{ij}(t) = \frac{G \pm |F_i - F_j|}{t_{ij}} \tag{9}$$

The ants search from node i to node j; if $F_i \geq F_j$, formula takes a positive sign. If $F_i < F_j$, formula takes a negative sign.

With the introduction of gravity pigment on inspired factor correction, making ant search progress has directivity, which can improve the search efficiency of the whole ant colony algorithm.
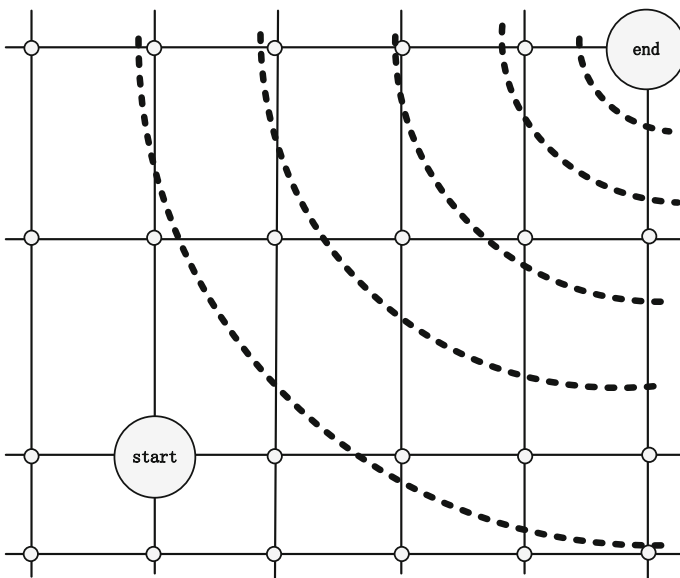

Fig. 4 Influence schematic of gravitational pigment

Due to the positive feedback principle in ant colony algorithm, it does not only accelerate the speed of the ant in the search, but also could be trapped in local optimal route search. In this paper, the theory of chaos is used; by adding the chaos disturbance, the random nature of the selection probability is increased, and the problem of falling into local optimal problem is overcome.

Motion randomness generally obtained by deterministic equations is called chaos. Logistic mapping is a typical chaotic system, and iterative formula is:

$$X_{i+1} = \mu \cdot X_i \cdot (1 - X_i), \quad \mu \in (2, 4] \tag{10}$$

where μ is the control parameter. When $\mu = 4$, $0 \le X_i \le 1$, the equation is completely caught into chaotic state. In this paper, the transition probability is improved as follows:

$$P_{ij}^k(t) = \begin{cases} \dfrac{\tau_{ij}^{\alpha}(t)\eta_{ij}^{\beta}(t)\left(1+x_{ij}\right)^{\gamma}(t)}{\sum_{s\in\text{allowed}_k}\tau_{is}^{\alpha}(t)\eta_{is}^{\beta}(t)\left(1+x_{ij}\right)^{\gamma}(t)}, & j \in \text{allowed}_k \\ 0, & \text{others} \end{cases} \tag{11}$$

where x_{ij} is chaotic variable, and γ represents the relative importance of chaotic perturbations.

Shared private parking finalization

After the initial selection of private shared lots, combined with the various factors of quantitative indicators, building weighted TOPSIS model ultimately determines the optimal private lot [14].

TOPSIS stands for Technique for Order Preference by Similarity to Ideal Solution, weighted different value based on the importance degree of each evaluation index, for finding the optimal solution and the worst solution in the evaluation process. The distance between each evaluation object and the optimal solution and the worst solution is calculated, respectively, and the relative approach degree of each evaluation object to the optimal solution is obtained, as the basis to evaluate the quality of each object [15]. Based on weighted TOPSIS, comprehensive evaluation analysis of the specific steps is as follows:

1. **Evaluation matrix preprocessing**
(1) Determining the indicators and specific parameter values

First of all, parking users choose their consideration indicators in the parking process. $E = \{e_1; e_2; \ldots; e_n\}$ is the indicator set, such as travel time, parking cost and walking distance. And then parking users determine the specific parameter values.

(2) Establishing the original data matrix

Set $P = \{P_1, P_2, \ldots, P_m\}$ as an alternative, which represents the collection of parking spaces. $E = \{e_1, e_2, \ldots, e_n\}$ is the indicator set, which represents parking choice evaluation index. We can then establish original data matrix $A = \left(x_{ij}\right)_{m \times n}$.

(3) Establishing the same trend Matrix

In private shared lot evaluation index, not all the indicators are the same direction. For example, the value of parking safety and parking convenience is the bigger the better. However, the value of walking distance, travel time, and parking cost is the smaller the better. We need to carry on the same trend treatment to the evaluation index, so that all the evaluation indexes follow the same trend. The same trend process is as follows:

$$
y_{ij} = \begin{cases} \frac{x_{ij} - \min_i x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}, & i = 1, 2, \ldots, m; \ j \in E_1 \\ \frac{\max_i x_{ij} - x_{ij}}{\max_i x_{ij} - \min_i x_{ij}}, & i = 1, 2, \ldots, m; \ j \in E_2 \end{cases} \tag{12}
$$

where E_1 is benefit evaluation index, such as parking safety and parking convenience. E_2 is cost evaluation index, such as walking distance, travel time, and parking cost.

After the trend of transformation, the evaluation matrix becomes $Y = \left(y_{ij}\right)_{m \times n}$.

(4) Normalization

In order to eliminate the influence of different indicators between classes, normalization of matrix Y, and establish a standardized decision-making matrix Z.

$$
z_{ij} = \frac{y_{ij}}{\sqrt{\sum_{i=1}^{n} y_{ij}^2}} \quad i = 1, 2, \ldots, m \ j = 1, 2, \ldots, n. \tag{13}
$$

2. **Determining the weight of evaluation index**

In this paper, we use the weighting method of subjective and objective combination. $w_j = (w_1, w_2, \ldots, w_n)^{\mathrm{T}}$ is subjective weight, and $w_j^* = \left(w_1^*, w_2^*, \ldots, w_n^*\right)^{\mathrm{T}}$ is objective weight. Sets a and b are important combination coefficient, and the combination weighting method is calculated as follows:

$$
W = aw_j + bw_j^* \tag{14}
$$

where w is combination weight. Important coefficients a and b can be calculated as follows:

$$
\begin{cases} a = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} \left(x_j^+ - x_{ij}\right) w_j}{\sum_{i=1}^{m} \sum_{j=1}^{n} \left(x_j^+ - x_{ij}\right) \left(w_j + w_j^*\right)} \\ b = \frac{\sum_{i=1}^{m} \sum_{j=1}^{n} \left(x_j^+ - x_{ij}\right) w_j^*}{\sum_{i=1}^{m} \sum_{j=1}^{n} \left(x_j^+ - x_{ij}\right) \left(w_j + w_j^*\right)} \end{cases} \tag{15}
$$

x_j^+ is the optimal value of the index attribute in alternative set.

(1) Subjective weight

Subjective weighting method uses the method of AHP, by two evaluation indexes contrast to each other, in order to sort the importance of each index. Subjective weight is calculated as follows:

$$w_j = \frac{\sum_{i=1}^{n} r_{ij}}{\sum_{i=1}^{n} \sum_{j=1}^{n} r_{ij}} \tag{16}$$

where r_{ij} represents the important level of jth evaluation index in ith private lot.

(2) Objective weight

In general, if the difference in a certain attribute value is small among private shared lots, it shows that the effect of the evaluation index for shared lot selection is small. In particular, if the evaluation index of multiple attribute values is equal, the influence of the evaluation index for private shared lot selection can be ignored. From the point of view of decision making, no matter how important the evaluation index itself is, the deviation of the attribute value is the key to determining the importance of the evaluation index. The deviation is defined as follows:

$$v_{ij}(w) = \sum_{k=1}^{m} |z_{ij} - z_{kj}| w_j, \quad i = 1, 2, \ldots, m; \quad j = 1, 2, \ldots, n \tag{17}$$

$v_{ij}(w)$ represents the toll deviation, and weighted vector w should make total deviation maximum; set Lagrange function:

$$F(w, \lambda) = \sum_{j=1}^{n} \sum_{i=1}^{m} \sum_{k=1}^{m} |z_{ij} - z_{kj}| w_j + \lambda \left(\sum_{j=1}^{n} w_j^2 - 1 \right) \tag{18}$$

Partial derivative of w_j and λ and units of w, we eventually get the objective weight formula:

$$w_j^* = \frac{\sum_{i=1}^{m} \sum_{k=1}^{m} |z_{ij} - z_{kj}|}{\sum_{j=1}^{n} \sum_{i=1}^{m} \sum_{k=1}^{m} |z_{ij} - z_{kj}|} \tag{19}$$

3. **Decision ranking**

(1) Calculate the normalized weighting matrix

$$U = \left(u_{ij}\right)_{m \times n} = \left(w_j z_{ij}\right)_{m \times n} = \begin{bmatrix} u_{11} & u_{12} & \cdots & u_{1n} \\ u_{21} & u_{22} & \cdots & u_{2n} \\ \vdots & \vdots & \ddots & \vdots \\ u_{m1} & u_{m1} & \cdots & u_{mn} \end{bmatrix} \qquad (20)$$

where w_j is the weight of index j.

(2) Obtaining the optimal value vector and the worst value vector

Based on the weighted standardized matrix row vector $U_i = \left(u_{i1}, u_{i2}, \ldots, u_{in}\right)^{\mathrm{T}}$, to determine the positive ideal point u^+ and negative ideal point u^-.

$$u^+ = \left(u_{i1}^+, u_{i2}^+, \ldots, u_{in}^+\right)^{\mathrm{T}}, \quad u_j^+ = \max\left[u_{ij}\right], \quad j = 1, 2, \ldots, m \qquad (21)$$

$$u^- = \left(u_{i1}^-, u_{i2}^-, \ldots, u_{in}^-\right)^{\mathrm{T}}, \quad u_j^- = \min\left[u_{ij}\right], \quad j = 1, 2, \ldots, m \qquad (22)$$

(3) Calculating Euclidean distance and proximity

Using the Euclidean norm distribution to calculate the Euclidean distance of the optimal solution D_i^+ and the worst solution D_i^-:

$$D_i^+ = \sqrt{\sum_{j=1}^{n} \left(u_{ij}^+ - u_{ij}\right)^2} \quad i = 1, 2, \ldots, n. \qquad (23)$$

$$D_i^- = \sqrt{\sum_{j=1}^{n} \left(u_{ij}^- - u_{ij}\right)^2} \quad i = 1, 2, \ldots, n. \qquad (24)$$

The relative approach degree of each evaluation object and the optimal solution C_i:

$$C_i = \frac{D_i^-}{D_i^- + D_i^+} \quad i = 1, 2, \ldots, n. \qquad (25)$$

Application

In this paper, we select the residential area of Sino-Japanese Friendship Hospital of Jilin University as the application site. If a parking user drives from the Nanling campus of Jilin University east gate to Sino-Japanese Friendship Hospital of Jilin University, departure time is at 10:30 a.m. and is expected to park for one hour. The evaluation indexes of the users' selection of parking lot mainly include travel time, walking distance, parking cost, parking safety, and parking convenience.

First of all, conduct the preliminary selection of the private shared lot around the Sino-Japanese Friendship Hospital of Jilin University, the position of the Sino-Japanese Friendship Hospital of Jilin University as the center and the maximum walking distance of 500m as the radius to search for a suitable private shared lot within the area. As a result, 5 possible private shared lots are identified, as shown in Fig. 5.

According to the road travel time calculation method, we determine the weight of each segment based on travel time, using improved ant colony algorithm, and determine the best route from the Nanling campus of Jilin University east gate to the five alternative private shared lots, respectively. The driving optimal route is shown in Fig. 6.

Fig. 5 500-m walking distance range coverage area

Fig. 6 Starting point to the alternative shared lot optimal path

Table 3 Attribute value of each evaluation index of private shared lot

Lot number	Lot type	Travel time (min)	Walking distance (m)	Parking fee (yuan)	Parking safety	Parking convenience
1	Open-air lot	14	350	4	4	9
2	Underground lot	13	200	5	8	6
3	Open-air lot	15	385	3	5	7
4	Underground lot	11	453	2	9	4
5	Residential underground lot	15	475	2	6	8

Table 4 The relative proximity to the optimal solution of private shared parking

The relative proximity degree	C_1	C_2	C_3	C_4	C_5
	0.33	0.54	0.24	0.65	0.26

Fig. 7 Optimal routes to private shared lot

After determining the travel time of each shared lot, combined with each sharing lot's factors, such as walking distance, parking fee, lot type, parking safety and parking convenience, to determine the evaluation index value. As shown in Table 3:

According to the weighted TOPSIS model, the relative proximity degree to the optimal solution for each alternative private shared lot is shown in Table 4.

By the relative proximity value, we can find number 4 lot is the optimal private shared lot, as shown in Fig. 7. The number 4 lot locates in No. 6 residential area, travel time is about 11 min, and parking cost is 2 yuan, after walking 453 m to reach the Sino-Japanese Friendship Hospital of Jilin University.

Conclusion

This paper puts forward a parking choice model based on private parking sharing. Firstly, we analyze and quantify the factors by parking users when selecting the private shared lot. Secondly, based on the maximum walking distance and shared-time period, the preliminary selection of private shared lot is conducted. Thirdly, improved ant colony algorithm is used to calculate travel optimal route and travel time from starting point to each alternative private shared lot. The weighted TOPSIS model is applied to calculate the optimal private shared lot. And last, the validity of the model is verified through application.

References

1. Qin, Yanping, Canqi Liu, and Ming Huang. 2008. Application of lot sharing in solving regional traffic congestion. *Urban Public Utility* 22 (3): 51–53.
2. Smith, Marry S. 1984. *Shared parking*. USA: Urban Land Institute.
3. Pan, Tingting. 2012. *Research on the opening up of residential parking space based on the "SPAT" theory*. Central South University.
4. Mou, Zhenhua. 2007. *Research on shared parking strategy of urban central areas under the compact city theory*. Huazhong University of Science and Technology.
5. Li, Fei. 2012. *The strategies for the mode of shared parking resource in settlements*. Dalian University of Technology.
6. Li, Chao. 2014. *City business circle parking characteristics and parking choice research*. Chongqing Jiaotong University.
7. Yu, Chengqiang. 2009. *A study on the user-oriented parking guidance and information system*. Fujian Agriculture and Forestry University.
8. Chen, Jun, Kai Xie, et al. 2015. Dynamic allocation model and effect evaluation of parking space sharing in central city university. *China Journal of Highway and Transport*.
9. Ji, Yanjie, Wei Wang, and Wei Deng. 2008. A method of parking guidance system in the parking guidance system before the trip. *Journal of Wuhan University of Technology (Transportation Science and Engineering)*.
10. Song, Meilin. 2014. *Research in optimal path model of intelligent parking guidance system*. Shanghai Jiao Tong University.
11. Dorigo, M., and L.M. Gambardella. 1997. Ant colony system: A cooperative learning approach to the traveling salesman problem. *IEEE Transactions on Evolutionary Computation* 1 (1): 53–56.
12. Colorni, A., M. Dorigo, and V. Maniezzo. 1991. *Distributed Optimization by Ant Colonies Proceedings. 1st European Conference on Artificial Life*. Paris, France: Elsevier, pp. 134–142.
13. Botee, H.M., and E. Bonabeau. 1998. Evolving ant colony optimization. *Advances in Complex Systems* 1 (2): 149–159.
14. Zhao, Shuchen. 2015. *Research on intelligent parking lot parking prediction and guidance method*. Harbin University of Commerce.
15. Sun, Lianchao. 2012. *Study on the choice model of city parking before the user oriented travel*. Chongqing Jiaotong University.

Mining Method for Road Traffic Network Synchronization Control Area

Lili Zheng, Hu Liu, Tongqiang Ding, Ruru Xing and Xinyue Hu

Abstract To ensure the road network synchronization control effect, it needs the correct division of synchronous control area. Therefore, according to the topology and traffic flow characteristics of road traffic network, this paper constructs road network modularity based on weights which are defined by the correlation degree between the intersections and designs synchronization control zoning method with the goal of maximizing modularity. At the same time, this paper uses road topology, traffic flow data, and tidal characteristics of the morning and evening peak time of business district of Guilin Road in Changchun and apples MATLAB and VISSIM to verify the analysis. The results show that the method is feasible and effective.

Keywords Road traffic network modularity · The key node · Condensation algorithm

Introduction

Urban road traffic system has the characteristics of discreteness, strong coupling, and nonlinear. For such a complex network, it is vital to prevent and control the cascading failure of system. In order to prevent the cascading failure of road transport network and massive traffic jams caused by the paralysis of key nodes, it needs to carry out synchronous traction control between the key nodes and other nodes within the traffic cells. Thus, the synchronization between intersections in

L. Zheng
State Key Lab of Automobile Simulation and Control, Jilin University,
Changchun 130022, China

L. Zheng · H. Liu · T. Ding (✉) · R. Xing · X. Hu
College of Transportation, Jilin University, Changchun 130022, China
e-mail: dtq8@163.com

L. Zheng
Jilin Province Key Laboratory of Road Traffic, Jilin University,
Changchun 130022, China

each district and the stability of the system can be achieved. But the key question is how to correctly divide synchronous control area and dynamic traffic control area.

At present, domestic and foreign researchers have made some achievements in terms of dynamic traffic zone division. Researchers generally carried out dynamic traffic zone division by building these principles that the adjacent intersection signal cycle and saturation are similar, and the distance is short and the degree of relevance is big. The methods covered cluster analysis, fuzzy reasoning, neural network method, and so on [1–9]. These studies obtained traffic zones from the local relationship between adjacent nodes. But dynamic traffic areas, especially synchronization control traffic zones, are a system. In the system, each node is correlated, inter-constraint and interactional, and these nodes form an organic whole. So we need to consider the overall performance of the system for dynamic traffic control zone division.

This paper applies complex network theory and considers the traffic flow characteristics to build road network modularity from the perspective of system optimal performance. Then, with maximizing modularity as the principle, it can acquire synchronization control area of the key nodes.

Road Traffic Network Modularity Construction

Consider an urban traffic network as an undirected weighted network $G = (V, E)$, where $V = \{v_1, v_2, \ldots, v_n\}$ is the set of all intersections and v_i is intersection i. $\lambda = \{\lambda_1, \lambda_2, \ldots, \lambda_n\}$ is the set of node importance index and λ_i is node importance index of the intersection i. $D = \{D_1, D_2, \ldots, D_n\}$ is node degree of the vector. $E = \{e_{12}, e_{23}, \ldots, e_{ij}\}$ is the set of the edges between two adjacent intersections, e_{ij} is the link between intersections i and j; the weight w_{ij} of edge e_{ij} is the correlation degree between intersections i and j. The connection matrix A is $n \times n$ dimensional symmetric matrix which means connection relationship between nodes. If two nodes are linked together, the corresponding element is 1, otherwise, the corresponding element is 0. The matrix B is edge weight matrix; if it has link between two nodes, then $B_{ij} = w_{ij}$, otherwise, the corresponding element is 0.

Weight Calculation

Traffic correlation degree reflects the characteristics of the traffic connection between links or nodes and describes the degree of correlation between links or nodes by considering travel time and road traffic flow as the parameter variables. The bigger the value of traffic correlation degree is, the greater the correlation degree between links or nodes is. Therefore, this paper regards traffic correlation degree as edge weight of traffic complex network:

Fig. 1 Connection type of two kinds of intersection

$$w_{ij} = \frac{0.5}{1+t} \left[\frac{n \cdot q_{\max}}{\sum_{i=1}^{n} q_i} - (N-2) \right] \tag{1}$$

where w_{ij} is the edge weight of traffic complex network, n is the number of branch which the traffic flow of upstream intersection flows into, for the cross intersection $n = 3$, $\sum_{i=1}^{n} q_i$ is the sum of traffic volumes which arrive at the downstream intersection, t is the travel time that vehicles running between the two intersection, and its unit is seconds, N is road lane number between upstream and downstream.

Due to the influence of road restrictions factors, there will be one-way and two-way roads appearing in the traffic network, as shown in Fig. 1. The weight of one-way road can be directly calculated by Eq. (1). The weight of two-way road is the maximum value of the weights of different directions.

$$w_{ij} = \max(w_{i \to j}, w_{j \to i}) \tag{2}$$

where $w_{i \to j}$ is the weight from node v_i to node v_j; $w_{i \to j}$ is the weight from node v_j to node v_i.

Road Traffic Network Modularity Modeling

Traffic network should not only consider the network topology, also consider the connection strength between nodes. The closer the connection between nodes is, the greater the weight is. In order to improve the accuracy of algorithm, this paper adds edge weight to improve the modularity by combining with the feature of road traffic as follows:

$$R = \frac{B_{in}}{B_{in} + B_{out}} = \frac{\sum_{i,j} B_{ij} \delta(i,j)}{\sum_{i,j} B_{ij}} \tag{3}$$

where B_{in} means the sum of edge weights which connect community internal nodes if the node belongs to this community. B_{out} means the sum of edge weights which connect community internal and external nodes. Assuming that the node v_i belongs to this community, then:

$$B_{ij} = \begin{cases} w_{ij}, & \text{node } i \text{ connected with node } j \\ 0, & \text{node } i \text{ not connected with node } j \end{cases} \tag{4}$$

$$\delta(i,j) = \begin{cases} 1, & \text{node } i \text{ and node } j \text{ belong to the community} \\ 0, & \text{node } i \text{ and node } j \text{ not belong to the community} \end{cases} \tag{5}$$

Partition of Synchronous Control Area based on Road Traffic Network Modularity

In order to ensure the implementation effect of the synchronization control, the algorithm adopted in this paper for partitioning is under the premise of identifying the key nodes. Centering on the key node, we constantly consolidate community along the direction of the road traffic network modularity increased to obtain the biggest modularity. In this way, the result of synchronous control area partition will be acquired.

The Identification of Key Node

The key nodes of synchronous control refer to that the nodes play a significant role in the global connectivity, network robustness, and efficiency exert, while their functions are lost. We use node importance degree index to evaluate the importance of nodes. Generally, the greater the node is, the more important the degree of index is. Overall, considering the two aspects of the traffic topology structure and the characteristic parameters, the node importance index λ relates to these factors: the node degree based on road grade, the average path length change rate ΔL, the betweenness based on travel time, the node flow, queue length, saturation, and delay, which contain seven influencing factors [10].

$$\lambda_\alpha = \sum_{j=1}^{7} w(j) r_\alpha(j) \tag{6}$$

$$r_\alpha(j) = \frac{\Delta_{\min} + \rho\Delta_{\max}}{\Delta_a(j) + \rho\Delta_{\max}} \tag{7}$$

$$\Delta_a(j) = \left| Y^*(j) - y_a^*(j) \right| \tag{8}$$

where λ is the node importance degree index; $w(j)$ is the corresponding weight affected by influencing factors j; $r_\alpha(j)$ is the correlation coefficient of the influencing factors j of node α; ρ is distinguish coefficient, and its value is between 0 and 1; $\Delta_{\min}(j) = \min\{\Delta_\alpha(j)\}$, $\Delta_{\max}(j) = \max\{\Delta_\alpha(j)\}$; $y_a^*(j)$ is the corresponding data of standardized influencing factors j; $Y_a^*(j)$ is the optimal value in the data $y_a^*(j)$.

Condensation Algorithm of Road Traffic Network Modularity Maximized

The main idea of the algorithm is as follows: First, assign node v_i which is directly connected with internal nodes of community C. The number of the nodes in community C is K. For each element node in the set H, assuming that it belongs to the community C, calculate the corresponding community modularity value R_{ci}. If $R_c < R_{ck} = \max\{R_{c1}, \ldots, R_{ci} \ldots, R_{cK}\}$, incorporate node v_k to community C, update community C and the variable set H, and order $R_c = R_{ck}$. Then search the variable set H again until $R_c > \max R_{ci}$, so that the community formed.

The search result is analyzed shown in Fig. 2c. Initially, key node is community C, as shown in Fig. 2a. Searching the nodes which are directly connected with corporate internal nodes of community C in the connection matrix A, these nodes are assigned to the variable set H, so $H = \{v_1, v_2, v_3, v_4\}$. Calculate R_{c1}, R_{c2}, R_{c3}, and R_{c4} separately. If $R_{c3} = \max\{R_{c1}, R_{c2}, R_{c3}, R_{c4}\}$, simultaneously $R_c < R_{c3}$, merge the node v_3 to community C and order $R_c = R_{c3}$, as shown in Fig. 2b. Update community C and the variable set H, then $H = \{v_1, v_2, v_4, v_5, v_6, v_7\}$. Modularity can be obtained that $R_{c5} = \max\{R_{c1}, R_{c2}, R_{c4}, R_{c5}, R_{c6}, R_{c7}\}$, and $R_c < R_{c5}$. Merge the node v_5 to community C and order $R_c = R_{c5}$. Update community C and the variable set $H = \{v_1, v_2, v_4, v_6, v_7, v_8, v_9\}$. Modularity can be obtained that $R_{c2} = \max\{R_{c1}, R_{c2}, R_{c4}, R_{c6}, R_{c7}, R_{c8}, R_{c9}\}$. At this time $R_c > R_{c2}$, community divided is over.
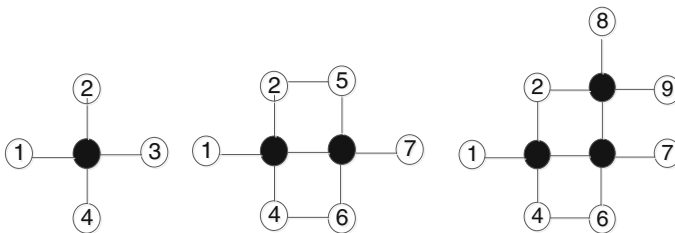
(a) Nodes searching   (b) Nodes mergence   (c) Result of Community partition

Fig. 2 Algorithm diagram

Steps of Synchronous Control Area Partition

The basic process of this method is shown in Fig. 3:

Step 1: Number intersections. Set up a corresponding traffic network connection matrix A. Obtain the node degree vector D.

Step 2: Calculating edge weight w_{ij} according to Eq. (1). Consider an urban traffic network as an undirected weighted network $G = (V, E)$. Establish the edge weight matrix B.

Step 3: Identifying the key node. Compute node importance index λ_i according to Eqs. (6)–(8). Select the largest λ_i as the key node G. Initially, the key node G is community C while the modularity $R_C = 0$.
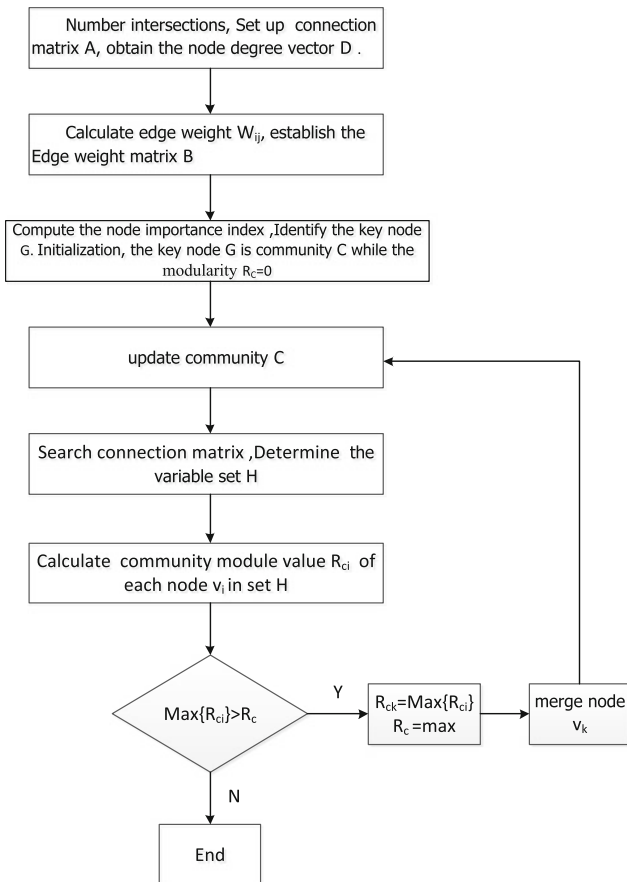
Fig. 3 Algorithm flow chart

Step 4: Searching the nodes v_i which is directly connected with internal nodes of community C. The number of the nodes in community C is K. For each element node in set H, assuming that v_i belongs to the community C, calculate the corresponding community the modularity R_{ci}. If $R_c < R_{ck} = \max\{R_{c1}, \ldots, R_{ci} \ldots, R_{cK}\}$, the node v_k should be incorporated to the community C. Update community C and the variable set H, order $R_c = R_{ck}$ and go to Step 5. Else if $R_c > R_{ck} = \max\{R_{c1}, \ldots, R_{ci} \ldots, R_{cK}\}$, go to Step 6.

Step 5: Repeat Step 4 and continue to merge node.

Step 6: At this time, the modularity gets to maximum. The algorithm is over. Acquire the synchronous control area based on key node.

Model Validation

This paper selects a local traffic network of Changchun city for empirical research surrounded by the Peony Street, Guilin Road, Fengshun Street, and Tsinghua Street, as shown in Fig. 4. To evaluate the effectiveness of the discussed method, simulation was ran to partition an urban traffic network under two different traffic scenarios which are off-rush hours and rush hours. Through surveying the import
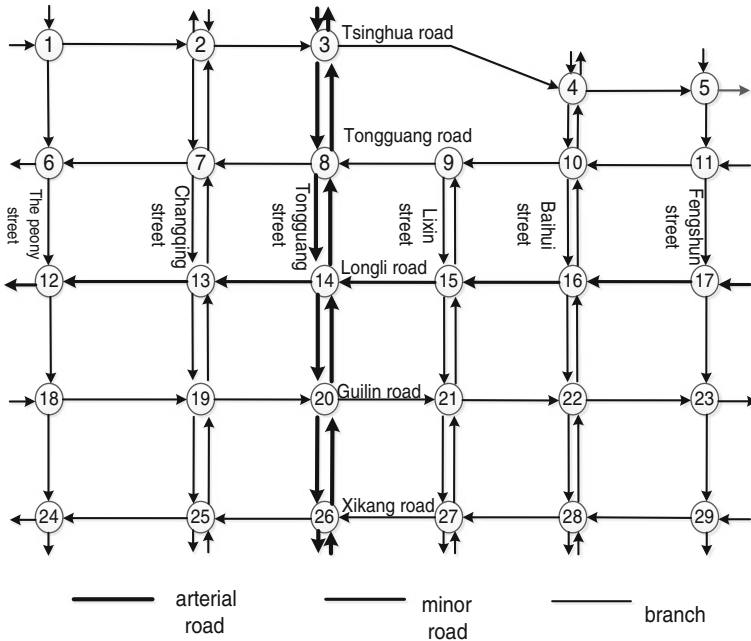
Fig. 4 Guilin road business district of Changchun urban road traffic network

lanes of each intersection, signal timing, import and export lane numbers, travel time and lane number, and other information, we use the simulation software VISSIM to simulate the different traffic network statuses of rush-hour traffic and off-rush hour traffic to obtain the data needed for the experiment. According to Eqs. (1) and (2), the network edge weights can be calculated during rush-hour traffic and off-rush hour traffic, as shown in Fig. 5a, b.

The traffic parameter data should be standardized by using MATLAB. Then, we can calculate each node importance degree index under rush-hour traffic status and off-rush hour traffic status according to Eqs. (6)–(8), as shown in Fig. 6.

Under rush-hour traffic status, the node importance degree index of node v_{14} is the biggest, whose value is 4.817. So node v_{14} is the key node. Search the nodes which are directly connected with internal nodes of community C in the connection matrix A and assign them to the variable set H, so $H = \{v_8, v_{13}, v_{20}, v_{15}\}$. While calculating the corresponding community modularity of each node in set H, we find that community module value of node v_{15} is the biggest. Hence, node v_{15} could be assigned to community C. Community C and the variable set H are updated and the search is continuing. When the community modularity is equal to 0.5617, it is no longer increased, as shown in Fig. 7. Therefore, synchronous control area $\{v_{14}, v_{15}, v_{16}, v_8, v_9, v_{10}\}$ with key node as the core has been got, as shown in Fig. 8.

(a) rush hour   **(b)** off-rush hour
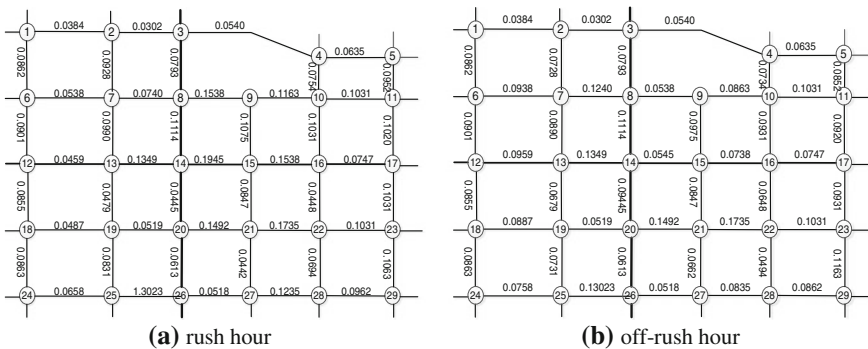
Fig. 5 Traffic network topology structure

Fig. 6 Node importance degree distribution under different traffic status
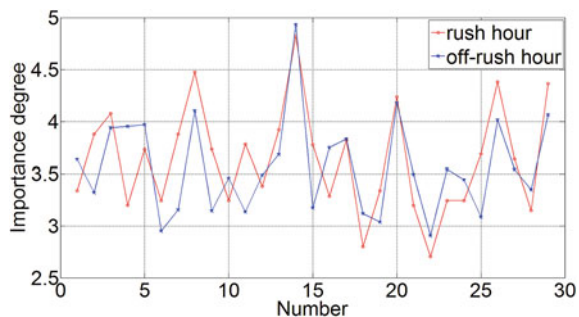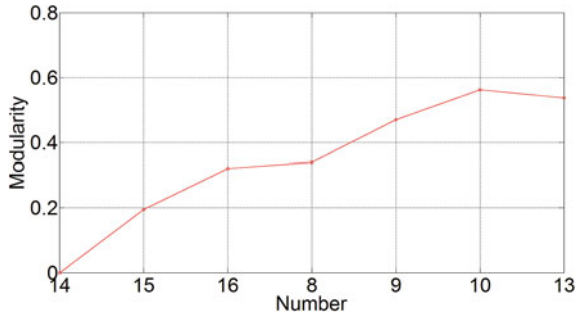
Fig. 7 Modularity curve during rush hour
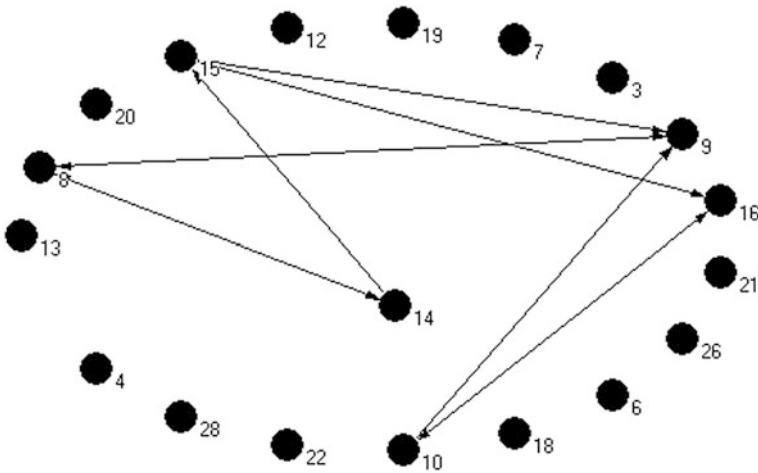
Fig. 8 Community combined figure during rush hour

In the same way, under off-rush hour traffic status, the node importance degree index of node v_{14} is the biggest, whose value is 4.932. So node v_{14} is the key node. Continuing to search and merge, when the community modularity of node v_{14} is equal to 4.932, the community modularity is no longer increased, as shown in Fig. 9. Therefore, synchronous control area $\{v_{14}, v_{13}, v_{20}, v_8, v_7, v_{19}\}$ is obtained with key node as the core, as shown in Fig. 10.

Compared with the two community divisions above, it can be known that under different statuses of rush hour and off-rush hour, node importance degree index and network edge weights can change, and community classification is also changed. Therefore, dynamic partitioning of synchronous control area can be realized.

The algorithm divides dynamic traffic area based on road traffic network modularity and selects module value maximum as the condition of the division end.
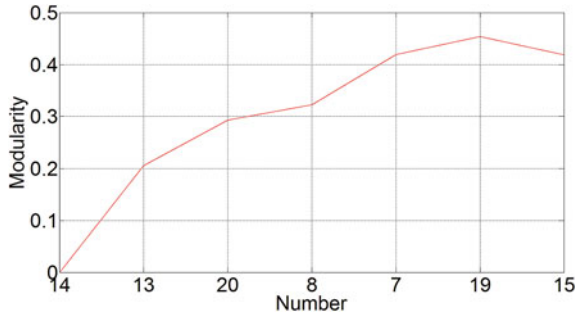
Fig. 9 Modularity curve during off-rush hour
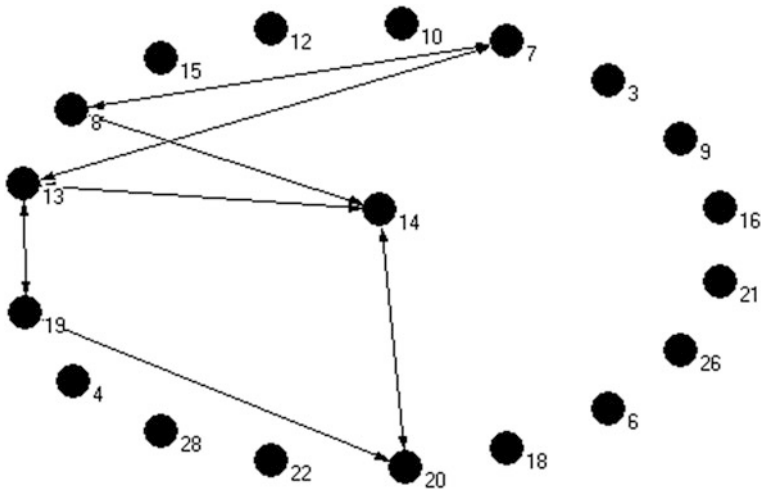
Fig. 10 Community combined figure during off-rush hour

The curve is shown in Figs. 7 and 9. Compared with other algorithms in traffic area division whose effect cannot be identified, this algorithm can divide dynamic traffic area while evaluating the overall performance of the community.

Conclusions

In this paper, an algorithm for urban community division based on key nodes is proposed, which is based on the module degree of complex network in the community and combines with the characteristics of urban road network and operation characteristics. And the algorithm is used in Changchun urban local road network for validation. The algorithm only needs to use network information and use the

evaluation index as the basis of community division. Thus, it improves the reliability of the corporate division and reduces the algorithm complexity.

According to the road traffic network topology and traffic flow characteristics, this paper applied complex network theory to build the road traffic network modularity. With the principle of maximizing the modularity, the method of synchronization control area division was designed. And the effectiveness of the proposed method is verified through using the actual road network. In view of the large-scale traffic network and to improve the operation efficiency of the algorithm, the author will discuss the corresponding parallel algorithm in the follow-up study.

References

1. Fan, Zhou, and Xiaofan Wang. 2007. A rapid local community partition algorithm of a complex network. *The Computer Simulation* 24 (11)
2. Yin, Hongying, and Liqun Xu. 2010. Urban road network dynamic partitioning based on spectral clustering algorithm. *Traffic Information and Security* 30 (12).
3. Li, Ruimin, and Huapu Lu. 2008. Traffic signal control area fuzzy dynamic partitioning method research. *Journal of Wuhan University of Technology (Transportation Science & Engineering)* 32 (3).
4. Hu, Yucong, and Haiwei Chen .2014. rban road network detection module structure and the Hub section diagnosis algorithm. *Journal of Southwest Jiaotong University*.49 (4).
5. Scott, J. 2002. Social network analysis: A handbook, 2nd ed. London: Sage Publications.
6. Liu, Shaohai, and Qingkun Liu. 2009. Complex network community method based on local module degree. *Computer Engineering and Design* 30 (20).
7. Clause, A. 2005. Finding local community structure in networks. *Physical Review E* 72: 026132.
8. Lin, L.T., and S.M. Tsao. 2002. A system approach on signal grouping for areawide control of computerized traffic system. TRB. In *Proceedings of the 79th TRB Annual Meeting*. Washington DC: TRB, 1–21.
9. Moore, J.E., and P. Jovanis. 1985. Statistical designation of traffic control subareas. *Journal of Transportation Engineering* 111 (3): 208–223.
10. Lili, Zheng and Shiguang Wang. 2015. Key node identification method of urban road traffic network. *Journal of Wuhan University of Technology (Transportation Science & Engineering)* 39 (4).

Research of Embedded Real-Time Passenger Flow Detection Equipment

Qing Tian, Liyuan Wang, Yun Wei, Xuece Pang, Qi Wang and Weiwei Fei

Abstract In order to further improve the accuracy and real-time performance of the passenger flow detection equipment, an improved passenger flow detection equipment based on embedded development platform was proposed. The equipment uses a new type of deep camera to gather the video image. We use the pedestrian detection technology based on the characteristics of head and shoulder on the embedded development platform to get real-time passenger flow information, and the technology based on the classical HOG algorithm, at the same time, the proposed technique can effectively improve the efficiency of the algorithm based on the depth information. Using network transport protocol implementation Ethernet high-speed transmission of data and video. The experiment shows that the equipment can achieve the real-time and accurate detection of the passenger flow in the complex environment. The average precision is 93.85% and the running time is 21 fps. Using the Ethernet interface, it has the advantages of easier development, lower cost, faster transmission speed, and remote control.

Keywords Passenger flow detection · Embedded development platform · Pedestrian detection · Network transmission

Q. Tian (✉) · L. Wang · X. Pang · Q. Wang
School of Electronic Information Engineering, North China University of Technology, Beijing 100144, China
e-mail: tianqing@ncut.edu.cn

Y. Wei
Beijing Urban Construction Design and Development Group Co. Ltd., Beijing 100037, China

W. Fei
Systems Engineering Research Institute CSSC, Beijing, China

Introduction

In recent years, urban rail transit is accepted by more and more people because of its large capacity, fast, and other characteristics. At the same time, the contradiction between the passenger flow and the carrying capacity is becoming increasingly significant, which leads to the problem of safety and efficiency. Timely and accurately accessing the passenger flow information can provide the necessary data to forecast the development trend, release early warning information for managers in the case of large passenger traffic. It can enhance the security risk identification ability of managers, and effectively reduce the rate and severity of the accidents.

Pedestrian detection algorithm is the core algorithm of passenger flow detection equipment, with the development of computer hardware and video capture equipment. Many typical pedestrian detection algorithms are proposed [1–3]. A pedestrian detection algorithm based on histogram of oriented gradient (HOG) and support vector machine (SVM) was proposed by Dalal et al. [2]. It has become one of the classical algorithms in the field of pedestrian detection. Felzenszwalb et al. [4] improved the HOG algorithm proposed by Dalal et al. and proposed a part-based algorithm based on the HOG algorithm. Tang et al. [5] proposed a method of using several cameras of different visual angles to detect passenger flow. In addition, target tracking is another important aspect of the passenger flow detection. For this purpose, domestic and foreign scholars have conducted a lot of research and exploration in the field of pedestrian tracking [6]. At present, common tracking algorithms are based on feature [7], 3D model [8], deformable template [9], and region [10], etc.

This paper combines and optimizes the classical algorithm. The classic pedestrian detection method based on HOG feature extraction has a large amount of calculation, especially on the embedded platform. It cannot meet the need of real-time detection. Because the depth image provides accurate distance information, contour and other notable features, the depth image is filtered by setting the height threshold method. It can get the human head and shoulder area. Within this region, it extracts HOG feature. This algorithm effectively reduces the computation of HOG significantly and improves the efficiency.

Algorithm Design of Passenger Flow Detection Equipment

Intelligent video analysis part is the core part of the passenger flow detection equipment, this part includes image acquisition, depth image pretreatment, target extraction of head and shoulder, pedestrian detection, and pedestrian trajectory tracking. The specific steps are shown in Fig. 1.

In the complex environment, change of pedestrian poses and mutual occlusion between pedestrians affects the results of pedestrian detection, which caused false and missed detection. For this problem, this paper presents a method. It uses the
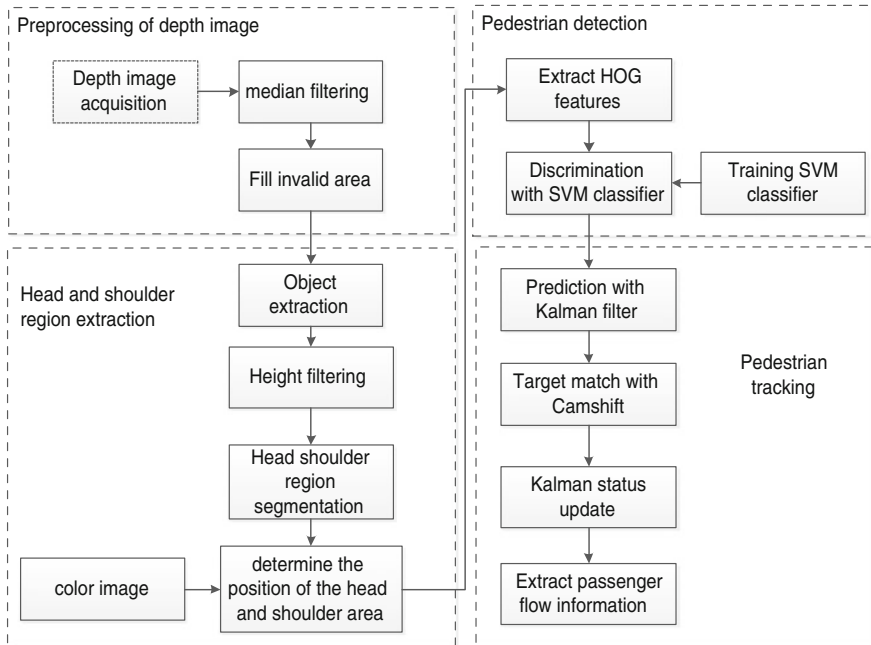
Fig. 1 Algorithm of passenger flow detection equipment

head and shoulder as the interest region, which is not easy to be blocked and the poses changed indistinctly.

Due to the specificity of the depth camera [11], after collecting the depth image, we initialize the image. First, we do median filter to the image in order to wipe off noise, at the same time, protect image's edge feature so as to make it unambiguous. Then we take the nearest nonzero value to replace an invalid area in the space domain. After the initialization of the image, it is easy to extract feature.

Then, using the background subtraction method for splitting foreground object, it can exclude the interference of objects in the scene to detect the object, extract detection objects, and narrow detection range. The depth value is normalized to 0–255; according to the angle of the camera installed in the configuration file and the trigonometric function theorem, we transform the mean of gray value from the linear distance between the pixel point and camera to pedestrian height. By setting the pedestrian height threshold to extract the head and shoulder image, filter out other areas, which narrow the scope of the next feature extraction.

We proposed an efficient extract objects feature method, the objects are classified by feature, in order to avoid the people's clothing, posture, and bags on the shoulder influence to detection results. This equipment uses pedestrian detection method of combining HOG and SVM to complete the above functions [12]. The specific steps of HOG are as follows:

(1) Global image standardization

Gamma correction method for global image standardization is used to reduce local shadow and exposure effects to HOG features.

The formula for gamma correction is:

$$I(x, y) = I(x, y)^{\text{gamma}} \tag{1}$$

(2) Image gradient calculation

Image gradient calculation can enhance the contour of the image. We can use one-dimensional discrete gradient operator, Laplace operator, and Sobel operator for image gradient calculation. In this paper, we use a simple and efficient one-dimensional discrete gradient operator. One-dimensional horizontal and vertical gradient operator are defined as $[-1, 0, 1]$ and $[-1, 0, 1]^T$, and the horizontal and vertical gradient of pixel point (x, y) are calculated as:

$$G_x(x, y) = I(x+1, y) - I(x-1, y) \tag{2}$$

$$G_y(x, y) = I(x, y+1) - I(x, y-1) \tag{3}$$

where I is image pixel value, G_x, and G_y are gradient in the horizontal and vertical direction of the pixels. Pixel point (x, y) gradient magnitude is:

$$G(x, y) = \sqrt{G_x(x, y)^2 + G_y(x, y)^2} \tag{4}$$

Gradient direction is:

$$\alpha(x, y) = \tan^{-1}\left(\frac{G_y(x, y)}{G_x(x, y)}\right) \tag{5}$$

(3) Unit gradient histogram construction

The image is divided into several units, constructing nine bins histograms for each unit to count the gradient information of the pixels in the unit.

(4) Block normalization

In the difference of contrast between foreground objects and background and the change of local illumination, which leads to a large change in the calculated gradient, it needs to be normalized gradient intensity. The normalized function is usually used with the following:

$$\text{L1}-\text{norm}: v \leftarrow v/(\|v\|_1 + \varepsilon) \tag{6}$$

$$L1-\text{sqrt}: v \leftarrow \sqrt{v/(\|v\|_1 + \varepsilon)} \tag{7}$$

$$L2-\text{norm}: v \leftarrow \sqrt{\|v\|_2^2 + \varepsilon^2} \tag{8}$$

In which v is not normalized histogram vector, $\|v\|_k$ is k order normal form of vector v, ε is a tiny constant, in order to prevent the divisor is 0.

(5) Generation of HOG feature vector

Combine all the features of the overlapping blocks together to form the final feature description vector.

After extracting HOG feature, using the SVM classifier for classification. The function of SVM classifier is to calculate the optimal segmentation of these features according to certain rules to classify and recognize objects. Then, target tracking will find the candidate target position similar to the target feature in the video image sequence by appropriate search matching algorithm and effective target features, so as to calculate the traffic flow information. This paper uses the way of combining CamShift and Kalman to track target. The CamShift algorithm can adjust the object's area when the adaptive objects size changed. It can continue to track the target. Kalman filter can effectively predict the object's position, avoid the influence of background, and target occlusion on the tracking results to some extent. After calculating the target trajectory, the method of setting count lines is used to calculate the two-way traffic flow in the scene.

The Structure and Hardware Design of the Passenger Flow Detection Equipment

Structure of Passenger Flow Detection Equipment

This passenger flow detection equipment is composed of three parts: the image acquisition unit, the video image processing unit, and the POE power supply unit. A new type of depth camera is used to complete the video capture function, using the USB to connect the image acquisition unit and the video image processing unit. After the analysis of video, the passenger flow information is transmitted to the server through an Ethernet controller, at the same time, the equipment transfer video to network video recorder. The collection and analysis unit is supplied by the POE power supply unit. Passenger flow detection equipment, servers, and network video recorder are connected with the switch to form the LAN, in order to greatly improve the speed of data transmission (Fig. 2).
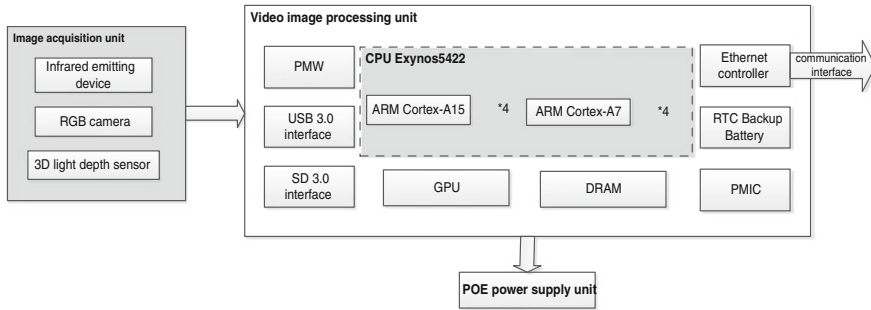
Fig. 2 Hardware structure of passenger flow detection equipment

Hardware Design of Passenger Flow Detection Equipment

The equipment uses a new type of deep camera to capture the color and depth images within sight range. Compared with traditional cameras, this new type of camera collects the color images and depth images in the scene at the same time. The camera is composed of three parts, which are infrared projector, color camera, and infrared camera. The maximum resolution is 640×480 and the frame rate can reach 30 fps.

Samsung Exynos5422 chip is the core part of the video image processing unit, it is equipped with four Cortex-A15 ARM and four Cortex-A7 ARM processors, and the highest frequency of the two parts are 2.1 and 1.5 GHz, respectively. According to the actual operation, the chip will assign an important calculation to the four Cortex-A15 core, it will allocate easy calculation to the four Cortex-A7 core. In this way, the real-time video image can be analyzed efficiently. The GPU used ARM Mali-T628 MP6, which is a kind of integrated image processing chip, and it can run 1–8 core according to the need. MP6 is a clock with six GPU units, its frequency is up to 533 or 600 MHz. Equipment using GPU can greatly improve the speed of image processing, and then improve the efficiency of video processing. Due to the good performance of GPU, the power consumption will be greatly improved, the heat is larger, and it needs fan cooling. The equipment uses a microprocessor PWM square wave to control the fan, and then based on the GPU temperature to determine whether or not to increase the fan speed. Ethernet controller use RTL8153 chip with the most advanced DSP technology and mixed signal technology to achieve the purpose of high-speed information transmission.

Lithium battery connects to the PMIC of the equipment, and I2C of the PMIC give RTC power supply to ensure that RTC is still running when external power is closed. POE power supply unit uses an Ethernet cable to ensure the transmission of data while providing power supply to equipment. This way greatly reduces the difficulty of the equipment installation wiring, and installation costs will be reduced as well.

Software Design of Passenger Flow Detection Equipment

The software design of the equipment includes intelligent video analysis unit, video transmission unit, and equipment management unit. Three units realize the video acquisition, analysis, data transmission, preservation, and equipment maintenance. This design ensures the accuracy and stability of the system work while improving the system's functions. It greatly reduces the cost of post maintenance and equipment management (Fig. 3).

Software Design of Intelligent Video Analysis Unit

The intelligent video analysis unit includes video capture module, parameter setting module, video analysis module, data collecting module, and data transmitting module. It collects real-time color images and depth images in the surveillance area by depth camera. Using the video processing technology, the equipment can extract the effective passenger flow information and sends the data to the server.

First, the equipment initializes operation environment, and then opens the camera, sets the image size, format, frame rate, and other related parameters, then it creates and opens the color image and the depth image data flow, starts the image acquisition. Equipment registers image-driven mode, synchronize color images and depth images, and read equipment installed height, angle, and other parameters in
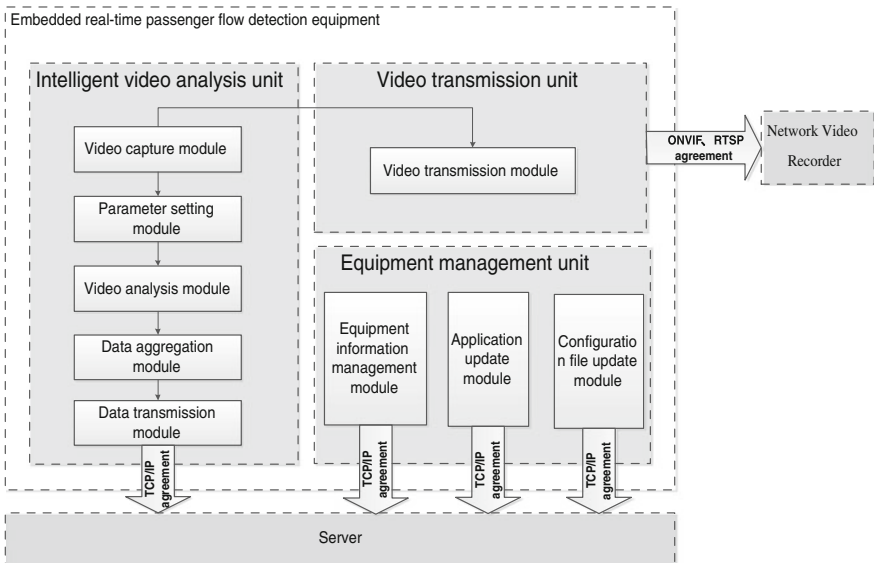
Fig. 3 Software structure of passenger flow detection equipment

the configuration file. Connecting equipment and the server sends equipment IP address, MAC address, and other basic information to the server, and equipment receives incoming data streams. According to the pedestrian detection and tracking algorithm to calculate the passenger flow of each frame. It will summary each frame of the passenger flow information, accumulated two-way passenger flow per 5 s. This data is packaged in a specified format to be sent to the server, when the server receives successfully, it will return an instruction, the equipment will clear the previous data, and begin processing the next frame image (Fig. 4).
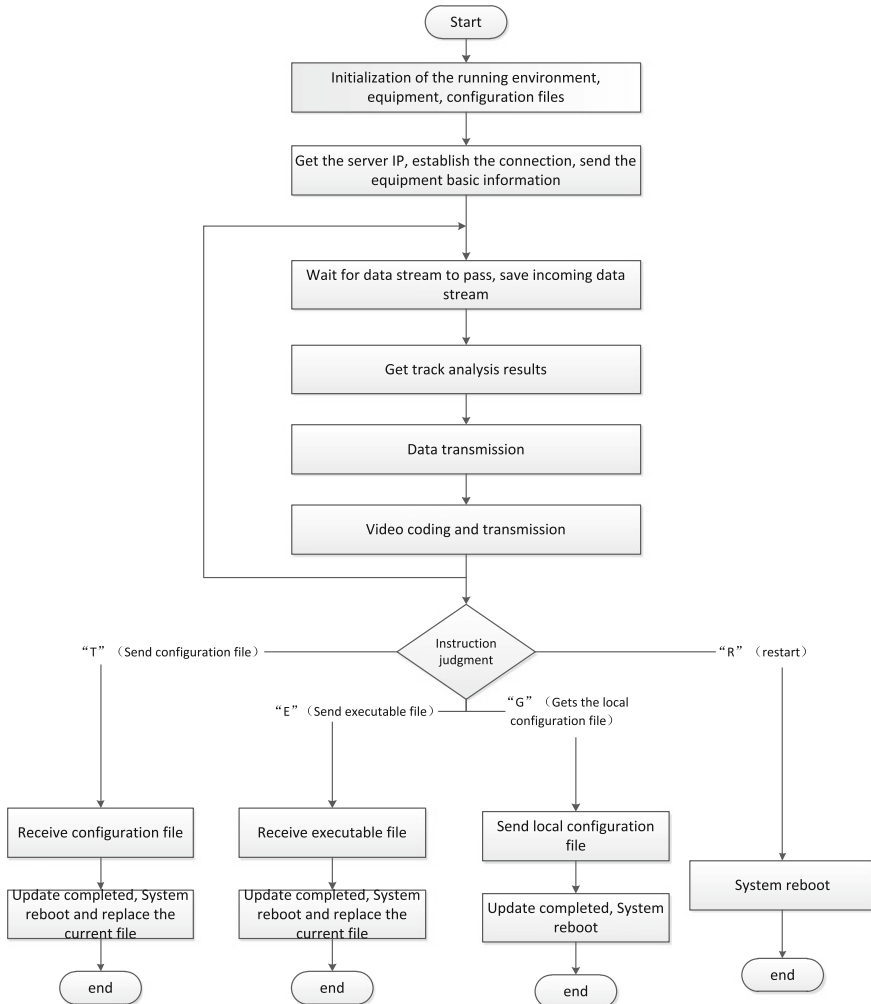
Fig. 4 Software flow chart of passenger flow detection equipment

Software Design of Equipment Management Unit

When it is necessary to update the equipment file, the server sends an instruction to suspend the transmission of data. The equipment stops sending data and waiting for receiving instruction. After receiving the instruction successfully, according to the instruction, the equipment begins to receive the executable file, configuration file, or send the local configuration file to server, after the completion of the operation of the file, the equipment is restarted, replaces the updated files, and recovers equipment operation.

Software Design of Video Transmission Unit

After camera capturing color images, equipment compresses and encodes the image. H264 frame data temporarily stores in the FIFO buffer queue, and it is transmitted to live 555 server, after live 555 server receiving the RTSP client requests, the FIFO queue reads H264 video data and makes a live video on RTSP. The equipment uses the ONVIF protocol, connects the network video recorder and the RSTP code stream, and then broadcasts the RTSP stream on the network video recorder.

Result of Experiment

The algorithm of this paper is completed on the platform of Ubuntu14.04 and Opencv2.4.9, the new depth camera is used to acquire the image. Video resolution is 320 × 240, and the language programming uses standard C/C++.

In order to verify the efficiency of the detection and tracking algorithm and effectiveness of coding algorithm, we simulate the subway scene in the corridor of the laboratory building. Equipment installed height is 2.33 m, farthest acquisition distance is 5.97 m, equipment installation angle is 45°. In the daytime, indoor experimental results are as follows (the results of the experiment are shown in Table 1) (Fig. 5).

Table 1 Simulation experiment result

	Actual passenger flow	Equipment detection of the number of passenger flow	The number of false detection and missed	Correct rate (%)	Running time (ms)
Inbound passenger flow	113	109	4	96.5	45.70
Outbound passenger flow	114	104	10	91.2	

(a) The 1040th frame color image

(b) The 1060th frame color image

(c) The 1040th frame depth image

(d) The 1060th frame depth image
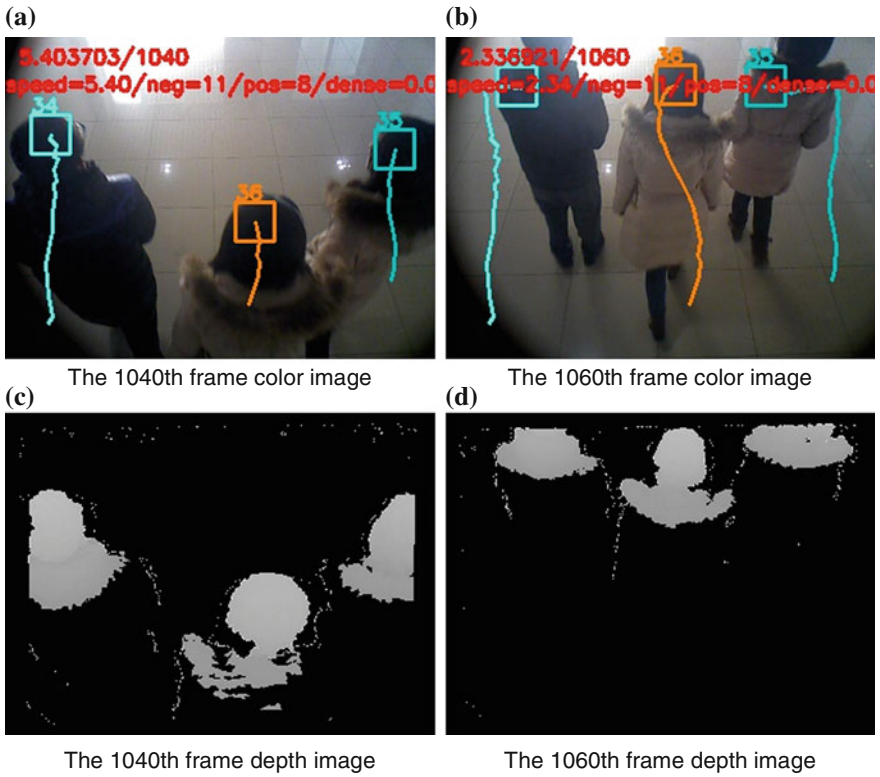
Fig. 5 Experimental result

From the chart, we can see that the human head and shoulder area can be obtained effectively by the equipment. The effect of pedestrian pose and occlusion on the detection result is effectively avoided. At the same time, the equipment can determine the position of the head and shoulder in the color image, which can track the trajectory of the pedestrian and count the passenger flow finally. Because the depth information is added in the algorithm, the computation of HOG feature extraction is greatly reduced, the efficiency of the algorithm is improved, and the equipment basically meets the needs of real-time video image detection.

In order to further verify the accuracy and real-time performance of the algorithm in this paper, we compared literature [14], which uses HOG and SVM algorithm for pedestrian detection in the color image, and literature [15], pedestrian detection algorithm based on the body contour features with our result of comparative accuracy and running time (a frame image average processing time). Results are shown in Table 2.

Compared with the experimental data, we found that the detection accuracy and running time of the algorithm is greatly improved. The accuracy rates are increased by 6.45 and 9.87%, respectively, and the running time reduced by 11.30 and

Table 2 Precision and running time comparison of three algorithms

	Average accuracy (%)	Running time (ms)	Image size
This paper	93.85	45.70	320 × 240
Literature [14]	87.4	57.00	352 × 288
Literature [15]	83.98	66.67	800 × 600

20.97 ms, respectively. In conclusion, the algorithm can ensure the accuracy of detection. At the same time, improve efficiency and shorten the running time.

Conclusion

This paper introduces a kind of passenger flow detection equipment. It includes pedestrian detection and tracking, data sending and receiving, video transmission, and equipment management functions. Equipment uses methods of combining HOG and SVM to achieve pedestrian detection, using the way of CamShift and Kalman combination to track the target. This algorithm can effectively obtain the passenger flow information, and the accuracy is higher. In this paper, based on the classic HOG algorithm, we add depth information to it, using the method of setting a high threshold, extract the human head shoulder effectively, reduce the amount of HOG feature extraction, and improve the efficiency of the algorithm. But because the video compression encoding efficiency is slightly lower, affecting the overall efficiency of the system. At the same time as the camera is unable to identify the body face or back to the camera accurately, the result of two-way passenger flow detection has a huge error. So this is the following issues which need to be improved.

References

1. Wei ,Y., Q. Tian, and T. Guo. 2013. An improved pedestrian detection algorithm integrating haar-like features and HOG descriptors. *Advances in Mechanical Engineering* (SCI).
2. Dalal, N, and B. Trriggs. 2005. Histograms of oriented gradients for human detection. In *IEEE Computer Society Conference on Computer Vision and Pattern Recognition, 2005 (CVPR 2005)*, vol 1, 886–893. IEEE.
3. Li, J.J., Q. Zhu, and Y.N. Wang. 2010. Detecting and tracking method of moving target in complex environment. *Chinese Journal of Scientific Instrument* 31(10): 2242–2247.
4. Felzenszwalb, P.F., R.B. Girshick, D. Mcallester et al. 2010. Object detection with discriminatively trained part-based models. In *IEEE Transactions on Pattern Analysis and Machine Intelligence* 32(9): 1627–1645.

5. Tang, C., Y.Y. Lin, M.F. Weng, and H.Y. Liao. 2015. Cross-camera knowledge transfer for multiview people counting. *IEEE Transactions on Image Processing* 24(1): 80–93.

6. Vidal, A, J.J. Marron, and M.A. Labrador. 2014. Real-time pedestrian tracking in indoor environments. In *IEEE Latin-America Conference on Communications (LATINCOM), 2014*, 1–6.

7. Lu, X.J., Z.R. Chen, T.H. Yin, and F. Yang. 2013. Research of pedestrian tracking based on HOG feature and haar feature. *Computer Science* 40(6): 199–203.

8. Kim, S., C.B. Park, S.W. Lee. 2006. Tracking 3D human body using particle filter in moving monocular camera. In *Proceedings of the 18th International Conference on Pattern Recognition*, vol 4, 805–808.

9. Kass, M., A. Witkin, and D. Terzopoulos. 1988. Snakes: Active counter models. *International Journal of Computer Vision* 1(4): 321–331.

10. Li, Q., C.F. Shao, and H. Yue. 2013. Pedestrian oriented multi-object tracking algorithm in video sequence.*Transactions of Beijing Institute of Technology* 22(2): 178–184.

11. Meng, M., F.B. Yang, Q.S. She, Y. Sun, and Z.Z. Luo. 2015. Human motion detection based on the depth image of Kinect. *Chinese Journal of Scientific Instrument* 36(2): 386–393.

12. Xu, C., M.Z. Gao, Y.F. Zha, and L.M. Cao. 2015. Bus passenger flow calculation algorithm based on HOG and SVM. *Chinese Journal of Scientific Instrument* 36(2): 446–452.

13. Li, F.S., Y.C. Zhang, H.J. Yang, and Y.P. Wang. 2014. Fast pedestrians counting algorithm based on HOG. *Computer Systems & Applications* 2014(5): 172–176.

14. Li, X.Y. 2013. *On the key research of the statistics of passengers on bus using multi-camera*, 26–53. Changchun: Jilin University.

15. Wei, Y., D.W. Li, G.F. Gao, and X.C. Zheng. 2015. Research on passenger crowd index for Beijing urban rail transit. *Urban Rapid Rail Transit* 28(3): 5–9.

Optimal Model of Timetable Under the Influence of Train Speed on the Utilization Rate of Regenerative Braking

Jia Feng, Xiamiao Li, Baohua Mao and Qi Xu

Abstract With expanded development of subway systems, regenerative braking energy utilization has been intensively studied from a timetable synchronize point of view. Yet, even successful modeling approaches such as those based on regenerative braking technique still do not fully take into account the fact that at lower than threshold speed, no energy can be utilized. To achieve a better performance on energy-efficient operation with the actual problem, this paper formulated a timetable rescheduling model (TR model) to optimize the timetable by maximizing the overlapping time between accelerating and regenerative braking actions of trains in the same substation with the consideration of regenerative braking process. We design a genetic algorithm (GA) to solve the model and present some numerical experiments based on the actual operation data of Beijing Yizhuang subway line in China. It is shown that the model can reduce the energy consumption by 25.3% compared with the current timetable. In addition, numerical experiments show that the dwell-time and lower limits of headway have a strong effort on regenerative energy utilization.

Keywords Subway system · Regenerative braking · Genetic algorithm

Introduction

With the rapidly growing economy and social development in China, the subway systems have expanded rapidly and the demand on energy consumption has escalated accordingly. The energy consumed by subway in China was equivalent to

J. Feng (✉) · X. Li
School of Traffic and Transportation Engineering, Central South University,
Changsha 410075, China
e-mail: fengjia@bjtu.edu.cn

B. Mao · Q. Xu
MOE Key Laboratory for Urban Transportation Complex Systems Theory
and Technology, Beijing Jiaotong University, Beijing 100044, China

a staggering amount of 141.98 million tons of standard coal in 2014 [1]. And the energy consumed by locomotives accounts for over 60% of the total operation consumption [2]. This huge energy cost and the environmental impacts make energy-saving operation in subway system an urgent issue. Traditional studies on energy-saving operation for subway systems mainly include mass reduction [3, 4], resistance reduction [5], and energy-efficient operation [6–9]. Recently, researchers have also paid more attention to the regenerative braking technique for the short distances between neighboring stations of most metro lines, usually within 1–2 km [10, 11].

Regenerative braking is an energy recovery mechanism which can convert the kinetic energy into electricity during the braking phase. The recovery energy generated by the braking train can be fed back into the overhead contact line and immediately be used by accelerating trains in the same electricity supply substation as shown in Fig. 1.

Although the utilization of regenerative braking energy is an important topic in the modern subway system, few results about the studies on regenerative braking energy utilization, especially regenerative braking technique, have been published in the literature. In 1972, Kokotovic and Singh [12] firstly proposed a locomotive operation strategy considering regenerative braking energy utilization. Gordon and Lehrer [13] developed a simulator of the train control to smooth out train trajectories through coordinated control of multiple trains and recovery from a backup. Albrecht [14] proposed an optimal model for effective use of regenerative braking energy by improving the interval between the departure times of two trains from the two different terminus stations (synchronization time). Ramos et al. [15] studied a timetable optimization model to maximize the overlapping time between accelerating and braking actions of trains in the same substation. Bae et al. [16] presented an analysis of regenerative energy and line energy consumption of two metro lines by testing the voltage and feeder currents of the lines. Miyatake and Ko [17] investigated an optimizing total energy consumption of multiple inverter train operation considering regenerative braking energy, and the results shown that train speed profiles and control inputs strongly depend on the phase of departure times. Malavasi et al. [18] described two models for estimating the energy-saving potential
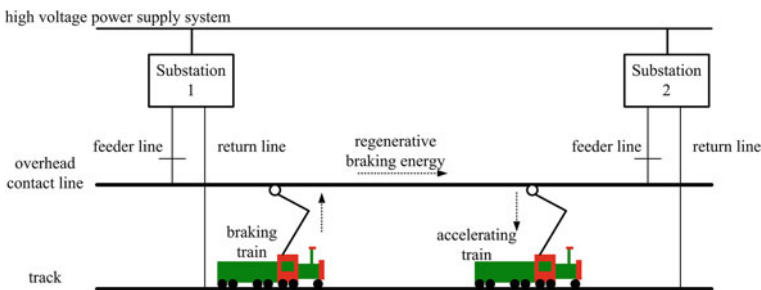

Fig. 1 Process of regenerative braking energy utilization

on a mass rapid transit system, but these models could not be achieved in actual system. Then, Peña-Alcaraz et al. [19] proposed a mathematical programming optimization model to synchronize the braking of trains arriving at station with the acceleration of trains exiting from the same or another station which improved robustness and regenerative braking energy utilization. Su et al. [20] proposed to optimize an integrated timetable which includes both the timetable and the speed profiles. A real-life Beijing Metro Yizhuang Line simulation study showed that the cooperative scheduling (CS) model can significantly reduce the entire route energy is 14.5%. Yang et al. [21] proposed a train CS rule to synchronize the accelerating and braking process of successive trains which can significantly improve the overlapping time around 22%. Based on the study [21], Li et al. [22] pointed that the run-time should be characterized as random variables for the departure delays for trains generally make the operation company reduce the run-time at following sections for minimizing the divergence from the planned timetable in some busy stations. They proposed a stochastic cooperative scheduling (SCS) approach. The results show that the SCS approach can save energy by 15.13% compared with the current timetable, and 8.81% compared with the CS approach. Su et al. [23] propose an integrated algorithm to generate the globally optimal operation schedule, and the computation time is within 2 s, which is short enough to be applied for real-time control system. Yang et al. [24] proposed a double-objective optimization model in order to minimize the energy consumption and minimizing passenger waiting time. Traction energy consumption is estimated based on the kinetic energy formula and the use of regenerative braking energy. Su et al. [25] proposed a cooperative train control model to minimize the practical energy consumption based on a common scenario that an accelerating train can reuse the regenerative energy from a braking train on the opposite track. According to the literature, the majority of the recent studies on regenerative braking technique are focused on timetable optimization for improving the utilization of regenerative braking energy by synchronizing the operations of accelerating and braking trains [9, 26]. Generally speaking, although few researchers considering regenerative energy utilization by maximizing the overlapping time between accelerating and braking actions of trains, most of them ignored the recovery rate decreases as the train speed decreases in real-life operation.

The theoretical and practice results show that when the speed of braking train is less than a certain value, the recovery process will be stopped, and the kinetic energy of the braking train will be converted to heat by friction and wasted. However, as one of the most lately raised topics in the regenerative braking technique, no clear definition can be found in the literature addressed this issue of recovery process will be stopped when braking train speed is less than the threshold value.

Taking more real-life factors of metro systems into consideration, we can know that the utilization of regenerative energy by coordinating control of multiple trains has been overstated. And the motivation of this paper is to clear how exaggerated
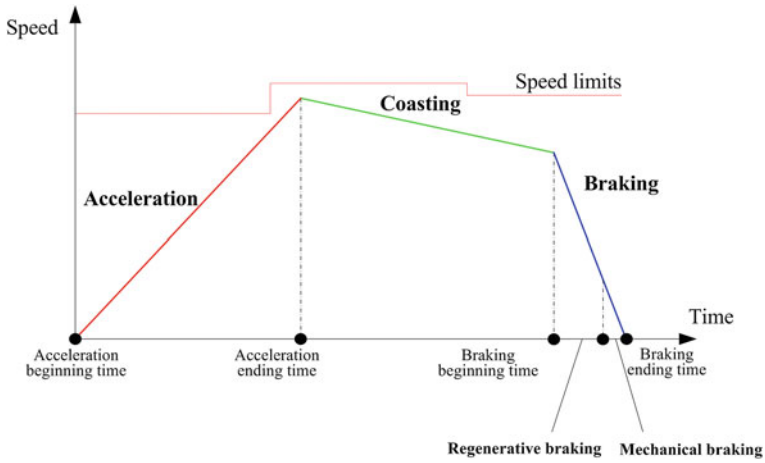
Fig. 2 Speed profile between successive stations

the utilization efficiency of regenerative energy have been. The proposed approach can be shown as follows:

(1) We revise the braking action of overlapping time counting as regenerative braking action, deleting mechanical braking action (see Fig. 2).

(2) We apply a parameter to describe how the train speed influenced the utilization of regenerative energy in the braking process.

The rest of this paper is organized as follows: in Sect. 2, the problem description and the overlapping time among successive trains definition are presented, then a timetable rescheduling model (TR model) considering the utilization efficiency of regenerative energy is proposed. In Sect. 3, the experimental results of the optimal timetables and the mechanism of regenerative braking energy utilization influence factors are discussed. Finally, conclusions are given in Sect. 4.

Problem Description and Modeling

The proposed model aims at improving the utilization of recovery energy by optimizing the timetable by maximizing the overlapping time between accelerating and braking actions of trains in the same substation with the considering of regenerative braking process.

A. **Problem description**

The recovery energy recovered from each braking train can be absorbed only by trains belonging to the same substation. To describe the optimal model simply, we make some assumptions as following.

(1) Assumption

    ① All trains running on the same direction share a common timetable except for the headway time, which means that they are assigned the same dwell-time on each station and the same run-time on each section.

    ② In all sections, the accelerating and braking times of trains are known parameters, which can be calculated by using the train's energy-efficient operation algorithm based on the current operation timetable.

(2) Parameters and variables

J	number of trains;
H_k	stations k;
K	number of stations H;
b	the braking deceleration;
T_k	travel time between station H_k and H_{k+1};
$T_{k,\,a}^{j}$	accelerating time between stations H_k and H_{k+1};
$T_{k,\,b}^{j}$	braking time between stations H_k and H_{k+1};
$t_{k,\,a}^{j}$	departure times of train j from station H_k;
$\mathbf{t_a}$	gather of departure times of train j from station H_k;
$t_{k,\,a}^{j+}$	ending time of the acceleration action of train j after departure from station H_k;
$t_{k+1,\,bj^-}$	beginning time of the braking action of train j arriving at station H_{k+1};
$t_{k+1,\,r}^{j}$	ending time of the regeneration braking action of train j arriving at station H_{k+1}, where $t_{k+1,\,r}^{j} = t_{k+1,\,b}^{j} - V_{\min}/b$;
$V_{\min}$	the threshold minimum speed for braking train action from regenerative braking to mechanical braking;
$t_{k+1,\,b}^{j}$	arrival time of the braking action of train j arriving at station H_{k+1};
$\mathbf{t_b}$	gather of arriving times of train j from station H_k;
$\varLambda$	the headway time;
D_k	the dwell-time required at station H_k;
R_k	the run-time of section between station H_k and station H_{k+1};
D_k^{l}	lower limit of dwell-time in station H_k;
D_k^{u}	upper limit of dwell-time in station H_k;
T^{l}	lower limit of total running time;
T^{u}	upper limit of total running time;
W	overlapping time of acceleration and braking actions of successive trains in the same substation;
$\lambda(k,\,k+1)$ a 0–1	variable, which takes value 1 if the two stations H_k and H_{k+1} belong to the same substation; otherwise, shall take value 0.
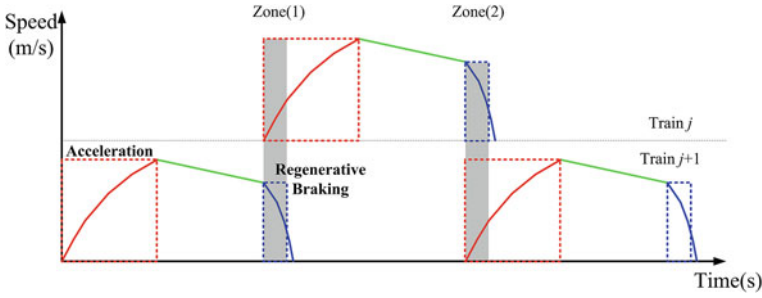
Fig. 3 Schematic of overlapping time calculation with different braking phase

(3) Operational rules

The regenerative braking energy can be better utilized for a smaller headway time during peak hours. The overlapping time calculation rule of the trains in succession is shown in Fig. 3.

As shown in Fig. 3, there are two kinds of situations of regenerative braking energy use when two trains run on the same direction. In Zone (1), train j accelerates to depart from station H_k, and train $j + 1$ decelerates to stop at station H_k. If both trains belong to the same substation, train j can absorb the regenerative braking energy from train $j + 1$. After headway time operation, train j decelerates to stop at station H_{k+1}, and train $j + 1$ accelerates to depart from station H_k. Then, train $j + 1$ can use the regenerative braking energy from train j which is shown in Zone (2).

B. Timetable optimization model

(1) Objective function

The goal of the optimization model is to maximize the overlapping time between the acceleration and braking actions of successive trains in the same substation. Take the Zone (1) for example, there are four scenes according to the run-time and dwell-time of successive trains, which can be expressed as Eq. (1) for each $1 \leq j \leq J$ and $1 \leq k \leq K$.

$$Q_1(\mathbf{t_a}, \mathbf{t_b}) = \begin{cases} 0, & \text{if } t_{k,a}^{j+} < t_{k,b}^{(j+1)-} \\ \min(T_{k,a}^j, t_{k,a}^{j+} - t_{k,b}^{(j+1)-}), & \text{if } t_{k,b}^{j-} \leq t_{k,b}^{(j+1)-} \leq t_{k,a}^{j+} \leq t_{k,r}^{j+1} \\ \min(T_{k-1,b}^{j+1} - t_{k,b}^{j+1} + t_{k,r}^{j+1}, t_{k,r}^{j+1} - t_{k,a}^{j}), & \text{if } t_{k,b}^{(j+1)-} \leq t_{k,a}^{j} \leq t_{k,r}^{j+1} \\ 0, & \text{if } t_{k,r}^{j+1} < t_{k,a}^{j} \end{cases}$$

(1)

In the same way, we can express the overlapping time of Zone (2) as follows

$$Q_2(t_a, t_b) = \begin{cases} 0, & \text{if } t_{k,a}^{(j+1)+} < t_{k+1,b}^{j-} \\ \min(T_{k,a}^{j+1}, t_{k,a}^{(j+1)+} - t_{k+1,b}^{j-}), & \text{if } t_{k+1,b}^{j-} \leq t_{k,a}^{(j+1)+} \leq t_{k+1,r}^{j} \\ \min(T_{k,b}^{j} - t_{k+1,b}^{j} + t_{k+1,r}^{j}, t_{k+1,r}^{j} - t_{k,a}^{j+1}), & \text{if } t_{k+1,b}^{j-} \leq t_{k,a}^{j+1} < t_{k+1,r}^{j} \\ 0, & \text{if } t_{k+1,r}^{j} < t_{k,a}^{j+1} \end{cases}$$

$$(2)$$

According to the relevant definitions of subway system operation, we define the decision variables as follows

$$\Lambda = t_{k,a}^{j+1} - t_{k,a}^{j} \tag{3}$$

$$D_k = t_{k,a}^{j} - t_{k,b}^{j} \tag{4}$$

$$R_k = t_{k+1,b}^{j} - t_{k,a}^{j} \tag{5}$$

Thus, the total overlapping time of all trains at all stations can be expressed as follows

$$Q(\Lambda, D) = \sum_{j=1}^{J-1} \sum_{k=1}^{K-1} [Q_1(\Lambda, D, j, k)\lambda(k-1, k) + Q_2(\Lambda, D, j, k)\lambda(k-1, k)] \tag{6}$$

(2) Constraints

In the operation of the subway system, the headway time, the dwell-time, and the run-time should be satisfied with certain restrictions, in order to keep system safety. To simplify the solution, we set the model for the integer programming, which means the variables are integer.

In the actual operation, dwell-time at each station has some flexibility in order to keep system robustness. We take ± 5 s and ± 10 s as the upper and lower limits of two different scenes.

Generally speaking, the plan operation time is often greater than the minimum operation time, and the difference between them is the train running recovery time, which is set for reliability of the train operation. Thus, we assume ± 5 and $\pm 10\%$ as the upper and lower limits of total running time.

(3) Optimization model

In conclusion, there is the optimization model as follows

$$\max Q(\Lambda, D) \tag{7}$$

$$\text{s.t. } D_k^l \leq D_k \leq D_k^u \tag{8}$$

$$T^l \leq \sum_{k=1}^{K-1} (D_k + T_k) \leq T^u \tag{9}$$

$$\Lambda, \ D_k \in Z, \ \ k = 1, 2, \cdots, K - 1 \tag{10}$$

C. Genetic algorithm

Genetic algorithm (GA) [27] is a stochastic search method for optimization problems based on the mechanics of natural selection and natural genetics, which was firstly initiated by Holland [28] in 1975 and applied to solve integer programming problems [29–33]. GA has strong global search ability, strong robustness, and high computation efficiency, so we apply GA to solve the proposed models.

The GA uses the principle of survival of the fittest to produce an approximate solution of the optimal solution in potential solutions. In every generation of GA, a new approximate solution will be created by fitness values of the individuals in the problem domain and choice method borrowed from nature genetics reconstruction until meets the convergence criteria.

For simply, we take binary coding to solve the model. A solution $(\Lambda, D_1, D_2, \ldots, D_{K-1})$ is represented by a chromosome $(C_1, C_2, \ldots, C_K)$, where gene C_i respects the ith decision variable in binary form. The basic procedure of the algorithm is presented as follows.

(1) Initialization: Randomly created chromosomes which number can meet the population quantity in the initialization.
(2) Calculate the fitness: We use the objective function in the model Eq. (8) as the evaluation function which is used to calculate the fitness of each chromosome. Then, take the maximum fitness chromosome as the best genes by arranging the chromosomes in the population according to the ascending order.
(3) Selection process: The selection of chromosomes is done by spinning the roulette wheel, which is a fitness-proportional selection. Each time, one chromosome is selected for a new child population.
(4) Crossover operation: Assume that the probability of crossover operation is P_c. Randomly select two parent chromosomes C_i and C_j and randomly generate a real number $r \in [0, 1]$. If $r < P_c$, C_i and C_j produce two new chromosomes X and Y through crossover operator. If X and Y are feasible, take them to replace their parents; otherwise, keep their parents.
(5) Mutation operation: Let P_m be the probability of mutation. Randomly select a chromosome C_i as the parent for mutation and randomly generate a real number $k \in [0, 1]$. If $k < P_m$, randomly select a gene of the chromosome and randomly select one bit of the gene. If the selected bit is 0 (or 1), take 1 (or 0) to replace it. If the new chromosome is feasible, take it to replace its parent; otherwise, keep its parent.

Experimental Studies

In order to illustrate the efficiency of the proposed model and solution method, we present some numerical experiments with different headway. The GA is coded in MATLAB 7.14 under the running environment: a Windows 7 platform of personal computer with processor speed 2.4 GHz and memory size 2 GB.

A. Calculating data

In this section, we present three numerical examples to illustrate the efficiency of the proposed model and solution method based on the data from Beijing Yizhuang subway line in China. This line is 23.23 km long with 14 stations and 6 substations. The current timetable is shown by Table 1, and the other parameters are listed as follows: $T_{k,a} = 25$ s, $T_{k,b} = 30$ s, $b = -1$ m/s^2, and $V_{\min} = 8$ m/s. When GA is performed, we set *pop size* = 200, $P_c = 0.8$, $P_m = 0.5$, and *max generation* = 100.

B. Example 1

In this example, we illustrate the efficiency of the integrated energy-efficient operation model. Take $\Lambda = 90$ s, which is used by most of the Beijing metro railway lines, the total overlapping time is 119 s, which is recorded in Table 2. For the original timetable, we take $\Lambda = 90$ s, and the total overlapping time is calculated to be 95 s, which implies that our approach can increase the total overlapping time by as many as $(119 - 95)/95 = 25.26\%$. The results are shown in Fig. 4. It illustrates that GA is convergent and *max generation* = 30 is a reasonable value.

C. Example 2

This example makes comparisons among CS model (CS, Yang et al. [22]) and TR model on regenerative energy utilization. Taking the headway is 90 s, the computation results are recorded by Table 1, where the sixth to eighth columns denote the optimal results of CS, the ninth to twelfth columns denote the optimal results of TR.

As can be seen, the overlapping time of optimal timetable by TR is 119 s, which is lower than 166 s counted by CS. However, these two overlapping times are not comparable. Their comparative criteria should be overlapping time of current timetable counted by TR and CS which are 95 and 136 s. It can be observed that TR has a better optimal ratio calculated by $(119 - 95)/95 = 25.3\%$ than CS (22.1%).

The reason for overlapping time counted by TR is lower than CS is obvious that the regenerative braking time is lower than braking time which is thought by CS.

D. Example 3

To illustrate the availability of GA, we set $T^l = T^u = 2087$ s and perform the proposed algorithm several times, with *pop size* = 200, $P_c = 0.8$, $P_m = 0.5$, and *max generation* = 30. The computation results are recorded by Table 2, where the second line denotes the overlapping time. The exact overlapping time calculated by

Table 1 Contrast of the optimal timetables with the same headway

Station	Substation	Current timetable			CS			TR		
		Arrival time	Dwell-time	Running time	Arrival time	Dwell-time	Running time	Arrival time	Dwell-time	Running time
1	I	0	30	190	0	25	190	0	26	190
2	I	220	30	108	215	25	108	216	35	108
3	I–II	358	30	157	348	25	157	359	25	157
4	II	545	30	135	530	35	135	541	35	135
5	II–III	710	30	95	700	30	90	711	32	90
6	III	835	30	114	820	35	114	833	31	114
7	III	979	30	103	969	35	103	978	32	103
8	III–IV	1112	30	104	1107	25	104	1113	30	104
9	IV	1246	30	164	1236	35	164	1247	34	164
10	IV–V	1440	30	150	1435	25	150	1445	25	150
11	V	1620	30	140	1610	35	140	1620	31	140
12	V–VI	1790	30	107	1785	30	102	1791	31	102
13	VI	1927	30	130	1917	50	115	1924	48	115
14	VI	2087	–	–	2082	–	–	2087	–	–
Overlapping time (s)		136 → **95**			166			*119*		
Optimal ratio		–			22.1%			*25.3%*		

Table 2 Best found values by GA

Time	1st	2nd	3rd	4th	5th	6th	7th	8th	9th	10th
Overlapping time (s)	119	118	119	120	119	118	119	115	119	117

Fig. 4 Optimization process of GA
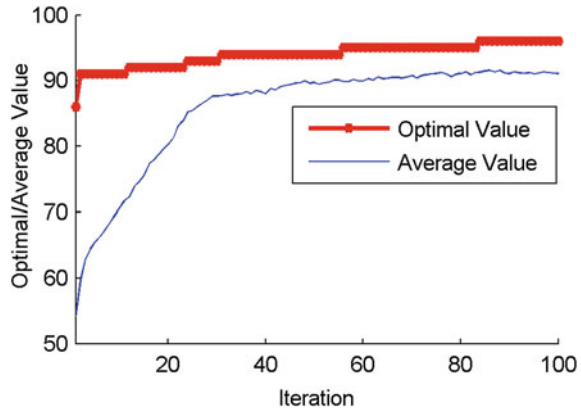
Matlab 7.14 is 119 s, and the average error is 0.59%, which demonstrates that GA is available.

E. **Example 4**

In order to explain the correlation mechanism of regenerative braking energy utilization influence factors (such as total running time, dwell-time, and headway), we analyze the relationship between regenerative braking energy and influence factors quantitatively by changing two factors and keeping one factor constant.

(1) **Analyze of total running time and headway**

A certain flexibility of total running time has a contribution to increasing the robustness of the timetable, improving the security and energy saving. Generally speaking, the running recovery time is about 10% of the total running time. So we set three scenarios: $T_{\text{limit}} = T \times (1 \pm 0\%)$, $T_{\text{limit}} = T \times (1 \pm 5\%)$, and $T_{\text{limit}} = T \times (1 \pm 10\%)$.

Taking the lower limit of headway from 78 to 98 s and the upper limit of headway from 180 to 360 s, the computation results are recorded by Table 3, where the second, fourth, sixth, and eighth lines denote the overlapping time, the third, fifth, seventh, and ninth lines denote the optimal headway.

It can be observed that the upper limits of the headway almost never affect the optimal headway. However, the lower limits of the headway have a strong influence on the optimal headway. When the lower limit is lower than 94 s, the optimal headway are equal to the lower limits. When the lower limit is larger than 96 s, the optimal headway are nearly equal to 120 s. Thus, we focus on the variation of the

Table 3 Overlapping time and optimal headway with different upper and lower limits of headway

Upper limits (s)	Lower limits (s)										
	78	80	82	84	86	88	90	92	94	96	98
360	211	193	177	159	138	114	96	82	72	70	70
	78	80	82	84	86	88	90	92	94	120	120
300	213	195	175	159	137	115	98	82	72	70	70
	78	80	82	84	86	88	90	92	94	118	119
240	209	193	175	156	139	115	96	82	72	70	70
	78	80	82	84	86	88	90	92	94	120	119
180	210	194	175	158	138	113	96	82	72	70	70
	78	80	82	84	86	88	90	92	94	117	118
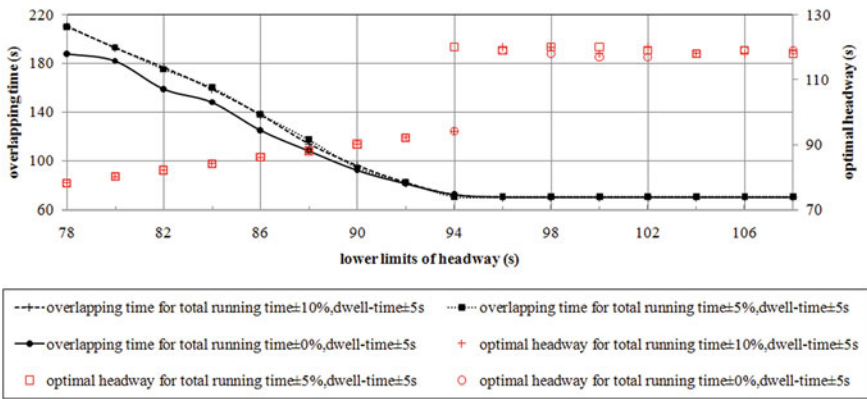
Fig. 5 Overlapping time and optimal headway with different total running time

overlapping time between different total running time and lower limits of the headway (see Fig. 5).

From Fig. 5, it can be seen that the three curves are very similar with a few difference. According to Fig. 6, it is easy to see that the overlapping time decreases sharply as the lower limits of headway increases from 78 to 94 s. However, the overlapping time is almost no change when the lower limit is greater than 96 s. Thus, the running time with flexibility makes a better effect than which with no fluctuation on energy saving.

However, the running time has hardly influences the optimal headway. According to the three types of points which indicate the optimal headway, we can see that the optimal headway is nearly equal to the lower limit when the limit is between 78 and 94 s and optimal headway is between 117 and 120 s when the limit is greater than 96 s.
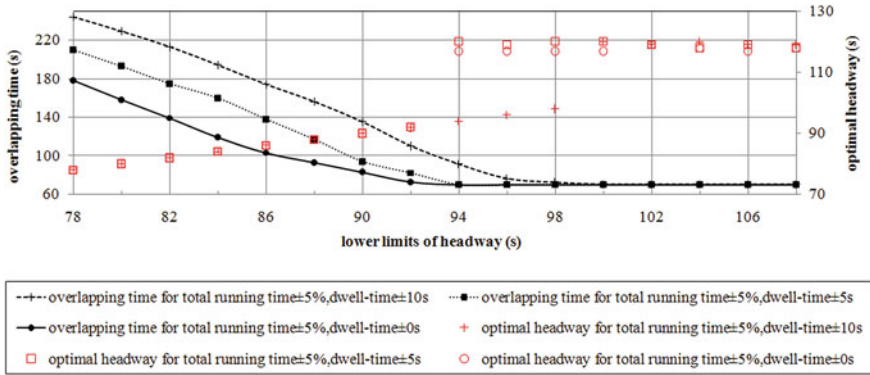
Fig. 6 Overlapping time and optimal headway with different dwell-time

(2) Analysis of dwell-time and headway

We set three scenarios about dwell-time with $T_{\text{limit}} = T \times (1 \pm 5\%)$: $D_{\text{limit}} = D \pm 0$ s, $D_{\text{limit}} = D \pm 5$ s, and $D_{\text{limit}} = D \pm 10$ s. Taking the lower limit of headway from 78 to 108 s and the upper limit of headway of 360 s, the computation results are recorded by Fig. 6.

As can be found in Fig. 6, the three curves are noticeably different when lower limit of headway increases from 78 to 94 s, but very similar when lower limit increases from 96 to 108 s. It is clear that the larger dwell-time range is, the more energy saving will be. Take lower limit of headway is 78 s for example, when dwell-time range is ±5 s, the overlapping time increases significantly from 178 to 210 s by 18.0% and when dwell-time range is ±10 s, the overlapping time increases significantly from 178 to 244 s by 37.1%. Thus, there is a clear possibility that the dwell-time is an influence factor of the total overlapping time. This is because the tracking interval of successive trains can be directly adjusted by the change of dwell-time, thus increasing the total overlapping time. Comparison of Figs. 5 and 6 shows that the dwell-time influences not only overlapping time, but also the optimal headway which is different from the running time.

Conclusions

Regenerative braking is useful for energy saving in subway system operation. An improved timetable rescheduling methodology to synchronize the accelerating and regenerative braking actions of trains in the same substation was proposed in this study which uses regenerative braking action instead of the braking action.

Extensive numerical experiments based on Beijing Yizhuang subway line were examined to demonstrate the effectiveness of the proposed approach. By comparing with CS model and TR model, we found that TR can save energy by 25.3%

compared with the current timetable. We clarified that CS only computes a larger overlapping time, but ignores the regenerative braking time. We also have examined the relationship between regenerative braking energy and influence factors (such as total running time, dwell-time, and headway), finding that the dwell-time and lower limits of headway have a strong effort on regenerative energy utilization.

The empirical findings provide insights regarding timetable rescheduling considering regenerative energy utilization. This solution method could be applied without difficulty to optimization of subway system operation. One of our future research directions is to extend to consider the variable acceleration rate, deceleration rate, and dwell-time which caused by stochastic passenger flow for making more realistic calculations of energy consumption.

Conflict of Interests The authors declare that there is no conflict of interests regarding the publication of this paper.

References

1. China Association Of Metros. Urban Rail Transit 2014 Annual Statistical Analysis Report. http://www.camet.org.cn/hyxw/201505/t20150513_407674.htm.
2. Feng, J. 2015. Train Behavior Optimization of Urban Rail Transit System Considering Energy Saving. PhD thesis, Beijing Jiaotong University.
3. IFEU (Institut fur Energieund Umweltforschung Heidelberg ¨GmbH). 2003. Energy Savings by Light-Weighting (Final Report), International Aluminium Institute (IAI), Heidelberg, Germany.
4. Chou, M., and X. Xia. 2007. Optimal cruise control of heavy-haul trains equipped with electronically controlled pneumatic brake systems. *Control Engineering Practice* 15 (5): 511–519.
5. Lukaszewicz, P. 2001. *Energy Consumption and Running Time for Trains: Modelling of Running Resistance and Driver Behavior Based on Full Scale Testing*. Stockholm, Sweden: Royal Institute of Technology.
6. Liu, R., and I.M. Golovitcher. 2003. Energy-efficient operation of rail vehicles. *Transportation Research Part A Policy & Practice* 37 (10): 917–932.
7. Bocharnikov, Y.V., A.M. Tobias, C. Roberts, S. Hillmansen, and C.J. Goodman. 2007. Optimal driving strategy for traction energy saving on DC suburban railways. *IET Electric Power Applications* 1 (5): 675–682.
8. Albrecht, A.R., P.G. Howlett, P.J. Pudney, and V. Xuan. 2013. Energy-efficient train control: From local convexity to global optimization and uniqueness. *Automatica* 49 (10): 3072–3078.
9. Li, X., and K.L. Hong. 2014. An energy-efficient scheduling and speed control approach for metro rail operations. *Transportation Research Part B Methodological* 64 (4): 73–89.

10. Liu, H.D., B.H. Mao, Y. Ding, W.Z. Jia, and S.K. Lai. 2007. Train energy-saving scheme with evaluation in urban mass transit systems. *Journal of Transportation Systems Engineering and Information Technology* 7 (5): 68–73.

11. Feng, X., B.H. Mao, X.J. Feng, and J. Feng. 2011. Study on the maximum operation speeds of metro trains for energy saving as well as transport efficiency improvement. *Energy* 36 (11): 6577–6582.

12. Kokotovic, P., and G. Singh. 1972. Minimum-energy control of a traction motor. *IEEE Transactions on Automatic Control* 17 (1): 92–95.

13. Gordon, S.P., and D.G. Lehrer 1998. Coordinated train control and energy management control strategies. Office of Scientific & Technical Information Technical Reports: 165–176.

14. Albrecht, T. 2004. Reducing power peaks and energy consumption in rail transit systems by simultaneous train running time control. *Power Supply, Energy Management and Catenary Problems* 3.

15. Ramos, A., M.T. Peña, A. Fernández, and P. Cucala. 2007. Mathematical programming approach to underground timetabling problem for maximizing time synchronization. International Conference on Industrial Engineering & Industrial Management—CIO, 1395–1405.

16. Bae, C, D, Jang, Y. Kim, S. Chang, and J. Mok, J 2007. Calculation of regenerative energy in DC 1500 V electric railway substations//7th International Conference on Power Electronics. ICPE'07. IEEE, 801–805.

17. Miyatake, M., and Ko, H. 2007. Numerical analyses of minimum energy operation of multiple trains under DC power feeding circuit//*2007 European Conference on Power Electronics and Applications*. IEEE, 1–10.

18. Malavasi, G., P. Palleschi, and S. Ricci. 2011. Driving and operation strategies for traction-energy saving in mass rapid transit systems. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 225 (5): 475–482.

19. Peña-Alcaraz, M., A. Fernández, A. Cucala, A. Ramos, and R. Pecharromán. 2012. Optimal underground timetable design based on power flow for maximizing the use of regenerative-braking energy. *Proceedings of the Institution of Mechanical Engineers, Part F: Journal of Rail and Rapid Transit* 226 (4): 397–408.

20. Su, S., Li, X., Tang, T., & Gao, Z. 2013. A subway train timetable optimization approach based on energy-efficient operation strategy. *IEEE Transactions on Intelligent Transportation Systems* 14 (2): 883–893.

21. Yang, X., X. Li, Z. Gao, H. Wang, and T. Tang. 2013. A cooperative scheduling model for timetable optimization in subway systems. *IEEE Transactions on Intelligent Transportation Systems* 14 (1): 438–447.

22. Li, X., and X. Yang. 2013. A stochastic timetable optimization model in subway systems. *International Journal of Uncertainty, Fuzziness and Knowledge-Based Systems* 21 (supp01): 1–15.

23. Su, S., T. Tang, X. Li, and Z. Gao. 2014. Optimization of multitrain operations in a subway system. *IEEE Transactions on Intelligent Transportation Systems* 15: 673–684.

24. Yang, X., B. Ning, X. Li, et al. 2014. A two-objective timetable optimization model in subway systems. *IEEE Transactions on Intelligent Transportation Systems* 15 (5): 1913–1921.

25. Su, S., T. Tang, and C. Roberts. 2015. A cooperative train control model for energy saving. *IEEE Transactions on Intelligent Transportation Systems* 16: 622–631.

26. Tuyttens, D., Fei, H.Y., Mezmaz, M., and Jalwan, J. 2013. Simulation-based genetic algorithm towards an energy-efficient railway traffic control. *Mathematical Problems in Engineering*, Article ID 805410, 1–12.

27. Arora, J.S. 2012. Genetic algorithms for optimum design. In *Introduction to Optimum Design*, 3rd ed, ed. Arora, J.S, pp. 643–655. Academic Press, Waltham, UK.

28. Holland, J.H. 1975. *Adaptation in Natural and Artificial Systems*. Ann Arbor: University of Michigan Press.

29. Lin, W.H., and C. Wang. 2004. An enhanced 0–1 mixed-integer LP formulation for traffic signal control. *IEEE Transactions on Intelligent Transportation Systems* 5 (4): 238–245.
30. Beard, C., and A. Ziliaskopoulos. 2006. System optimal signal optimization formulation. *Transportation Research Record Journal of the Transportation Research Board* 1978 (1): 102–112.
31. Pavlis, Y., and W.A. Recker. 2009. Mathematical logic approach for the transformation of the linear conditional piecewise functions of dispersion-and-store and cell transmission traffic flow models into linear mixed-integer form. *Transportation Science* 43 (1): 98–116.
32. Lertworawanich, P., M. Kuwahara, and M. Miska. 2011. A new multiobjective signal optimization for oversaturated networks. *IEEE Transactions on Intelligent Transportation Systems* 12 (4): 967–976.
33. Zhang, L., Y. Yin, and S. Chen. 2013. Robust signal timing optimization with environmental concerns. *Transportation Research Part C Emerging Technologies* 29 (1): 55–71.

Analysis of Spatial–Temporal Characteristics Based on Mobile Phone Data

Hong-liang Yin and Chang-jiang Zheng

Abstract The most traditional method to collect traffic data is household survey, which is a waste of manpower and material resources. OD matrices estimation using link volumes has been widely studied. But the significant shortcoming is the high cost of detectors. Besides, once installed in the network, the traffic detectors are not easy to be moved. Mobile phones' location data, however, can be acquired over a wider coverage without additional costs. The use of such data provides new spatiotemporal tools for improving urban transportation planning. This paper analyzes the nature and the pre-treatment of data from mobile phone operators in China and highlights the applicability of the data in domain of transportation. It also presents a typology of applications to analyze spatial–temporal characteristics based on mobile phone data.

Keywords Spatial–temporal characteristics · Mobile phone data · Urban transportation planning

Introduction

Recently, mobile communication technology is becoming one of the most widely used technologies, the most convenient and the most influential information and communication technologies (ICTs). The current penetration rates are 128 and 89% in developed and developing countries, respectively [1]. According to the Ministry of industry statistics of China, the number of mobile phone users in China has reached 1.256 billion by the end of May 2014 [2]. Along with the rapid growth of information techniques, the concept of smart city is becoming popular, which attracts the attention of urban planners from construction of cities and economy to

H. Yin
Jiangsu Provincial Department of Communications, Nanjing 210098, Jiangsu, China

C. Zheng (✉)
School of Civil Engineering and Transportation, Hohai University,
Nanjing 210001, Jiangsu, China
e-mail: zhenghhu@sina.com

citizens' activities. Hence, mobile phone users' footprints of their approximate locations, when making a call or sending an SMS, can show the daily routine that would interest urban planners.

Mobile phone data, as a new promising source of location data, consist of individual activities and are acquired almost in real time [3]. The processing, mining, publishing, and applying of such a huge amount of location data can provide the fundamental information of urban traffic demand, and therefore be a crucial for urban transportation planning of China during the rapid urbanization. It would be of great social and economic significance for cities to take advantage of those data and adopt new approach to ease traffic congestion, improve traffic safety and efficiency, reduce air pollution, etc. Data from mobile phone operators are becoming a tool for a smarter city in the future [4]. In recent years, mobile phone data have been used for estimating traffic flow [5, 6], deriving travel modes [7], analyzing interconnections between areas [8], tracking individual travel behavior [9], to name a few.

The main objective of the paper was to analyze how the use of mobile phone data can facilitate urban management and planning. The remainder of this paper is organized as follows: Section "Collection of Mobile Phone Data" is devoted to a description of the nature and the pre-treatment of data from mobile phone operators in China. Section "Analysis of Mobile Phone Data Applying in Urban Transportation Planning" discusses whether and how should such data be applied. Section "Key Technologies of Trip Characteristics Analysis Based on Mobile Phone Data" then gives a broad overview of the applicability of such data. Section "Conclusion" concludes the paper.

Collection of Mobile Phone Data

Interpretation of Mobile Phone Signaling

Original mobile phone data collected by the signaling acquisition system of China mobile consist of information such as service station number and location and switching time, [10]. Data format is shown in Table 1, and the detailed meaning of each field of mobile phone data is shown in Table 2.

Further more, MSID is the code of each mobile phone user given by mobile operators. LAC is usually divided by county or administrative district. CELLID is the code of base station where the mobile station locates. AREA CODE is the code of the region where the mobile station joins the network.

Mobile Phone Data Denoising

Traffic information collection based on mobile phone may be simple and efficient. The impact of communication network and the environment, however, adds much

Table 1 Original mobile phone data format of China mobile

MSID	TIMESTAMP	LAC	CELLID	EVENT ID	CAUSE	FLAG	MSCID	BSCID	CAUSE TYPE
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.30	13208	40482	8	9	1	11573	11483	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.40	13060	24403	8	9	1	11573	11584	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.40	13119	30062	8	9	1	11573	11512	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.40	13061	60331	8	9	1	11573	11563	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.10	13249	40573	8	9	1	11573	11540	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000000.90	33162	10391	8	9	1	11573	11405	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000000.90	33162	10392	8	9	1	11573	11405	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.30	13061	60621	8	9	1	11573	11563	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000000.90	13110	30291	8	9	1	11573	11750	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.60	13119	30582	8	9	1	11573	11512	4
c02c705e98588f724ca046ac59cafece65501e36	20130402000001.60	33141	31272	8	9	1	11573	11422	4
da4a1baac83a453050449d974ce0e1f50df1add3	20130402000002.40	13119	30473	14	9	0	11573	11512	4
a43ab53c7f4bc16c2de26e924a0d134f1103aef7d	20130402000000.70	13119	30141	5	9	0	11573	11512	4

Table 2 Name and implication of each field of the mobile phone data

Index	Field	No. of characters	Field meaning
1	MSID	32	Only identity code of user
2	TIMESTAMP	14	Signaling time, YYYYMMDDHH24MISS, unit: seconds
3	LAC	5	Location area (LA) code
4	CELLID	5	Zone number
5	EVENTID	3	The type of event, such as startup/shutdown, calling/called, normal location update
6	CAUSE	3	Cause of event
7	FLAG	3	Whether the identification can get IMSI
8	MSCID	8	ID of MSC
9	BSCID	8	ID of BSC
10	AREA CODE	8	Network area code

"noise" to the location data, which would influence the analysis. The noise usually results for the following two reasons [11]:

1. Switching disturbance (such as ping-pong effect), that is to say the frequent switches between two adjacent areas because of the judgment error of user's position switch in base station system. This kind of noise data is ubiquitous.
2. Staying at the same position during a long time. In a GSM network, a periodic location update function was introduced to prevent the loss of contact with a mobile station, which demands a mobile phone to report once every hour its position. Therefore, when the user stays at a fixed place, a large number of redundant data will be produced.

Those noise data are to the disadvantage for determining the track of mobile phone users. It is necessary to process them before analysis. For example, eliminate lines with abnormal value at the field of CELLID and LAC, due to communication failure, and lines with same values of MSID and CELLID, which are redundant data.

After pre-treatment, data need to be sorted by time. Mobile phone data of each user in each day should be extracted in order to calibrate their trajectories, where fields such as MSID, TIMESTAMP, and CELLID must be reserved.

Analysis of Mobile Phone Data Application in Urban Transportation Planning

Analysis of the Application of the Mobile Phone Data

Before applying the data, their application in transportation is analyzed as follows:

1. For traffic applications, the location and time information just fit for the traffic demand analysis. Besides, the characteristics of those data such as high sample rate, continuity, and real time give incomparable superiority to it.
2. The analysis of individual activity is to locate the traffic zone of each users and does not need the accurate location. Hence, the information of location and time satisfies the demand of analysis.
3. Object of data acquisition is mobile terminal, meanwhile most of the travelers in the city are cell phone users. The trip information analyzed from mobile phone data can therefore correctly reflect travel characteristics of habitants.
4. Though errors exist, the huge amount of data will fix the problem according to the theory of Big Data. The growing penetration rate guarantees the quantity of samples.

Research Framework of Urban Traffic Characteristics of Mobile Phone Data

The basic process of travel characteristics analysis based on mobile phone data is shown in Fig. 1. First, collect and extract information desired in the research area from mobile phone data platform of telecom operator. Second, establish the
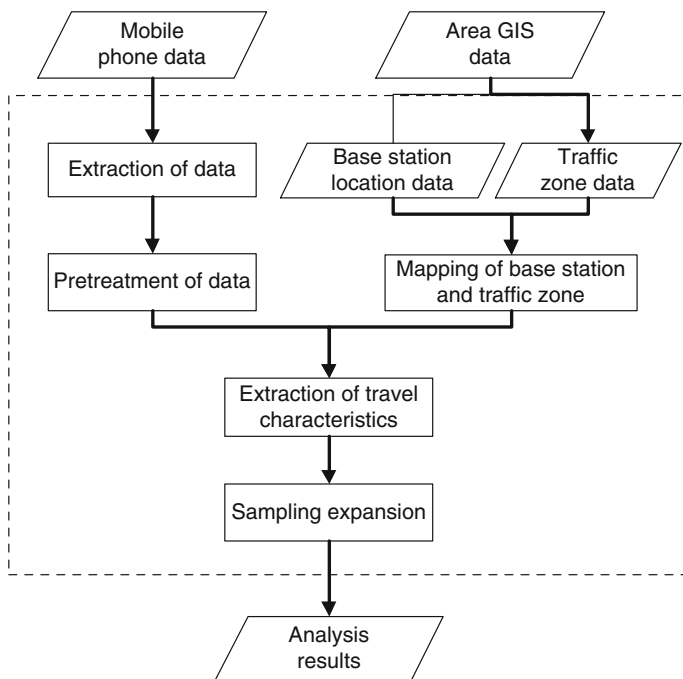
Fig. 1 Framework of travel characteristics analysis based on mobile phone data

mapping of base station and traffic zone data in GIS. Then, analyze travel characteristics with GIS data after pre-treatment of mobile phone data. Finally, get the analysis results by sampling expansion.

Research Area

During the extraction of travel characteristics from original mobile phone data, a mapping of base station zones and traffic zones need to be established. The analysis unit of mobile phone data is based on base station zones, which cover a certain area. The limit of the area, however, can be indefinite because of the instability of communication networks. There is, therefore, a deviation of user's location, which may range from 50 to 300 m in urban area and from 100 to 2000 m in Banlieue.

When the research area is divided into smaller traffic zones, the deviation of location data, especially in city center, will increase because of a greater number of base stations; while a traffic zone consists of base station span of control, this kind of deviation will decrease. In short, it will be more appropriate to apply mobile phone data to a grander range (like several cities) instead of a small area (like a few blocks).

Key Technologies of Trip Characteristics Analysis Based on Mobile Phone Data

Having huge advantages over traditional methods of traffic survey, extracting traffic information from mobile phone data is drawing more and more attention. Since 2005, several empirical studies have utilized data from mobile phone operators to gain insight and knowledge into complex and rapidly changing spatial urban phenomena. Research which utilizes the breadth of data produced by mobile phone operators can be a great opportunity to tackle specific issues. In this section, we will critically analyze several applications and focus on how to use such data.

OD Estimation

A general framework to develop OD matrices using mobile phone data is shown in Fig. 2.

First, generate the mapping relationship between traffic zones and base station zones according to the coverage situation of the telecom networks. Second, record the code of base station zones that a user has entered and left and the timestamp. Third, calculate the time that the user has stayed in every traffic zone. If the time exceeds a certain threshold, the zone can be seen as a stayed zone, while the others
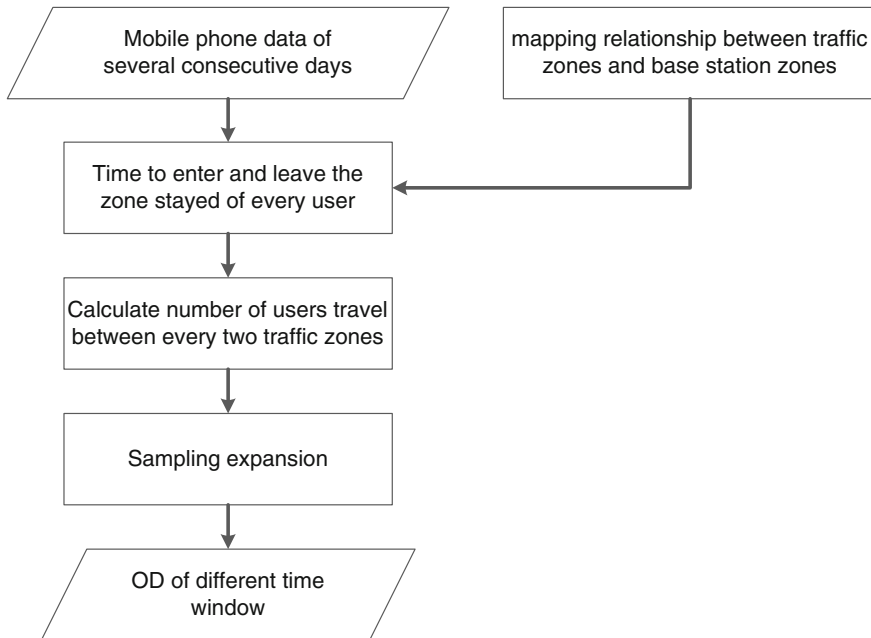
Fig. 2 OD estimation based on mobile phone data

are just passed through. Then, analyze trips occurring within certain time windows to generate transient OD matrices. Finally, considering the penetration rate of mobile phone users, get the OD estimation at different time by expanding samples to whole habitants.

Recognition of the Purpose of Trips

Taking the analysis of place of residence and work as example, the process is shown in Fig. 3.

Step 1: generate the mapping relationship between traffic zones and base station zones according to the coverage situation of the telecom networks.

Step 2: recognize the home zone of users, which should be the place where the user usually stays at night. For example, the zone where a user stays for more than 6 h between 20:00 and 8:00 and for more than 80% of the days studied is the user's home zone.

Step 3: recognize the working zone of users, which should be the place where the user usually stays during daytime. For example, the zone where the
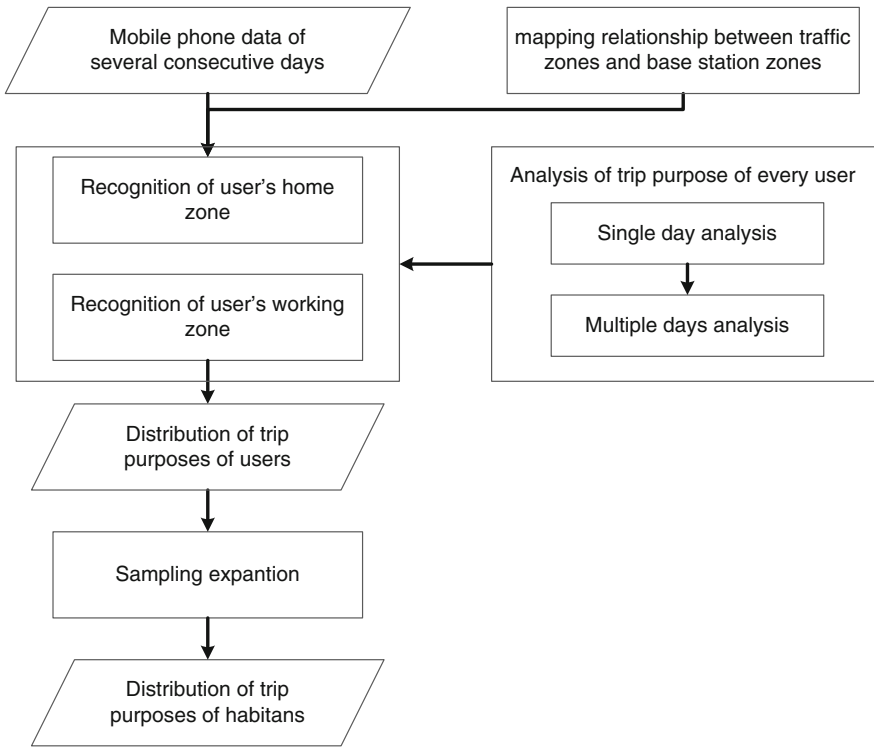
Fig. 3 Analysis of trip purpose based on mobile phone data

user stays for more than 6 h between 8:00 and 20:00 and for more than 80% of the days studied is the user's working zone.

Step 4: count the number of users that live or work in a certain zone to get the distribution of trip purpose of mobile phone users and expand the sample to obtain that of all the habitants.

Links Counts Detection

The process to detect traffic flow on links such as bridges and tunnels is shown in Fig. 4.

First, determine the link to analyze and establish the mapping relationship between telecommunication base stations and the two traffic zones on both sides of the link. Then, record the passing time if the trajectory of a user has passed the link. Finally, get the number of users who passed the link in different time windows.
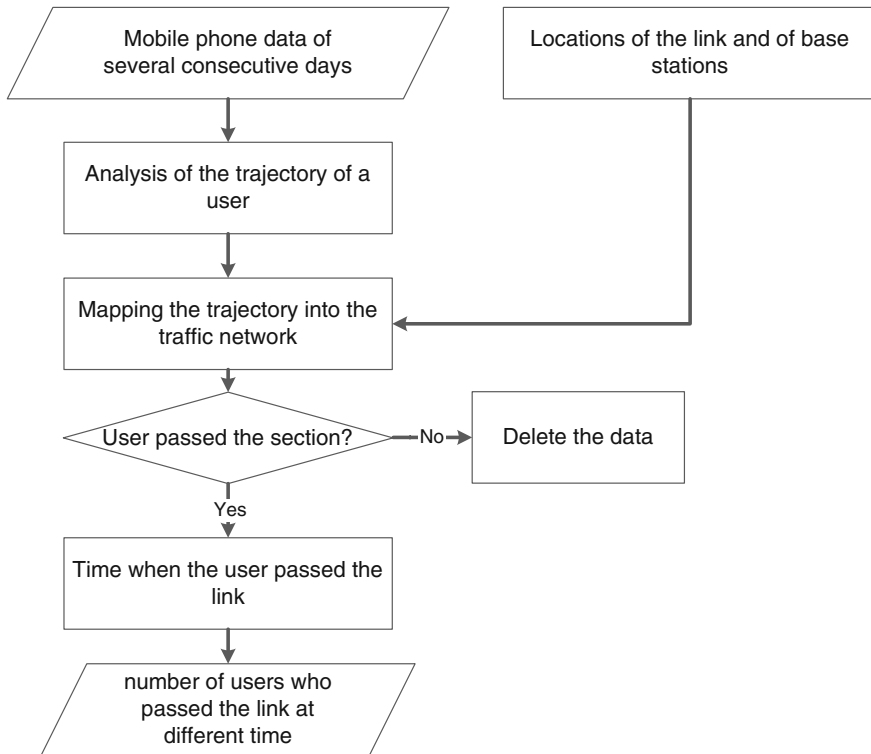
Fig. 4 Link counts detection based on mobile phone data

Conclusion

Better understanding of how, where, and when people move on a daily basis, especially in densely populated areas, could lead to improvement in urban planning, transportation infrastructure design, and assessments of environmental impacts [2, 11]. This paper has demonstrated the wide applicability of mobile phone data and the key technologies of application in achieving smart city objectives. Although such data cannot be approached as a panacea, it is hopeful that, in the near future, they would lead to a better city along with other sources of data.

References

1. Iqbal, M.S., C.F. Choudhury, P. Wang, et al. 2014. Development of origin–destination matrices using mobile phone call data. *Transportation Research Part C Emerging Technologies* 40 (1): 63–74.
2. Ran, B. 2013. Use of cellphone data in travel survey and transportation planning. *Urban Transport of China* 11 (1): 72–81.

3. Steenbruggen, J., E. Tranos, and P. Nijkamp. 2015. Data from mobile phone operators: A tool for smarter cities? *Research Memorandum* 39: 335–346.
4. Demissie, M.G., G.H.D.A. Correia, and C. Bento. 2013. Intelligent road traffic status detection system through cellular networks handover information: An exploratory study. *Transportation Research Part C Emerging Technologies* 32 (4): 76–88.
5. Wang, M., et al. 2013. Estimating dynamic origin-destination data and travel demand using cell phone network data. *International Journal of Intelligent Transportation Systems Research* 11 (2): 76–86.
6. Bohte, W., and K. Maat. 2009. Deriving and validating trip purposes and travel modes for multi-day GPS-based travel surveys: A large-scale application in the Netherlands. *Transportation Research Part C: Emerging Technologies* 17 (3): 285–297.
7. Trasarti, R., R. Trasarti, A.M. Olteanu-Raimond, M. Nanni, et al. 2014. Discovering urban and country dynamics from mobile phone data with spatial correlation patterns. *Telecommunications Policy* 39: 347–362.
8. Asakura, Y., and H. Eiji. 2004. Tracking survey for individual travel behaviour using mobile communication instruments. *Transportation Research Part C: Emerging Technologies* 12 (3): 273–291.
9. Guan, Z., X. Zhang, B. Hu et al. 2011. Research on decision support system of urban transport planning based on mobile phone data. In *Sixth Annual Conference of China Intelligent Transportation System*.
10. Huang, M., and B. Lu. 2010. Traffic OD data collection technology based on mobile phone location. *Journal of Chongqing Jiaotong University (Natural Science)* 29 (1): 162–166. (in Chinese).
11. Caceres, N., L.M. Romero, and F.G. Benitez. 2013. Inferring origin–destination trip matrices from aggregate volumes on groups of links: A case study using volumes inferred from mobile phone data. *Journal of Advanced Transportation* 47 (7): 650–666.

Analysis of Connection Mode and Type Between Regional Line and Urban Line of Urban Rail Transit

Guofei Gao and Yana Yan

Abstract As the main skeleton line, regional line is an important part of urban rail transit system and it will become a quick passageway between urban and peripheral towns such as airport, transport hub, and cluster. As a reference and guidance for the planning and designing staffs, this paper puts forward three main factors that are basic characteristics of the city, hierarchical system of urban rail transit system, and passenger flow factors. In addition, the influence way on the connection mode has been expounded. At the same time, three connection modes between the cohesion of the urban rail transit network and regional line have been put forward including existed types, adaptable condition in each mode choice of line selection, connection point, and other problems that should be paid attention to each connection mode.

Introduction

Regional line is an important part of urban rail transit system. With its unique advantages of fast and comfortable, it will become the main mode of transport in the long-distance transportation of urban rail transportation network. It not only extends the service scope of the urban rail transit which is given priority to the metro, but

G. Gao (✉)
School of Traffic & Transportation, Beijing Jiaotong University,
Beijing 100044, China
e-mail: gaoguofei_2007@sina.com

G. Gao
Beijing Urban Construction Design & Development Group CO.LTD,
Beijing 100037, China

Y. Yan
Urban Rail Transit Department, Beijing Vocational College of Transportation,
Beijing 100096, China

also enriches the content of the rail transit system. Regional line can improve travel speed and transferring convenience by connecting with existed rail lines and sharing tracks with existed rail lines; also, it achieves seamless connecting with the existed urban rail transit network through forming a large-scale transportation junction with outer rail transit system. Accordingly, not only can it anchor the urban rail transit network structure, but also can improve the efficiency of the whole urban rail transit network and optimize the overall structure of the network.

Names of regional line are not identical in major cities in developed countries, and the form is different as well. The regional line in Spain is called suburban line of national railway, In Southern California, it is called Metrolink; in New York, it is called commuter rail; in Canada, it is called AMT; in Japan, it is called JR; in France, it is called RER; and in Germany, it is called S-Bahn [1]. All of that are distributed in urban periphery and are systems-based medium and long distances. However, domestic regional line is still in research and at groping stage, which has not formed a system. There are a variety of forms that the line reaches into the city, such as arriving at the edge of the city, furthering, or running through the center of city. No matter which kind of form, there is a problem that how to connect between regional line and urban line. Whether the cohesion between each other is reasonable and efficient has practical significance to realize integration of regional and urban line. Therefore, in order to guide the benign development of the regional line, it is necessary to study transfer and connection between regional line and urban line of urban rail transit.

Analysis of Influencing Factors on Connection Mode

Influence of Basic Characteristics of the City on the Connection of Regional Line and Urban Line

Urban structure and passenger flow characteristics determine the network structure of rail transit, and at the same time, the construction process of rail transit also affects the evolution of the urban structure. The city with layers of structure, which always has a strong center and a number of subcenters or center groups, should adopt a ring radial network structure of rail transit. Based on the original loop line, rail transit network of Moscow has constructed several radioactive regional line that can link up with urban rail transit. As a coastal city, New York City has one seaside. At the same time, the Manhattan has absolute urban core status, so New York used the connecting way that regional line directly stretched into the center of the city and forms cohesion mode with multiple transfer spot. Generally, the size of domestic big cities is large, which has many new towns been planning and constructing, at the same time, rail transit construction is at early stage. If loop line connection mode is considered only between regional line and urban rail transit, there is a problem that the passenger flow of loop line in the line network will be larger. In Tianjin, the

comprehensive connection mode is adopted, and the regional line connected with the loop line and peripheral transportation hub in city fringe, it also it crossed the center of city [2]. Therefore, we should fully analyze urban structure and the characteristics of passenger flow, also we ought to plan rail transit network with the combination of a variety of connection mode between regional line and urban rail transit network.

Influence of Hierarchical System of Rail Transit System on the Connection of Regional Line and Urban Line

According to the spheres of city, city is divided into downtown, central city, and metropolitan area (including the suburban city circle and the outer suburban metropolitan circle). Land use, urban population, post, traffic demand, and travel strength are different in each circle. We should adopt corresponding means of transportation in each urban circle, according to the city circle division [3]. Subway that is a large capacity transportation is the solution of transportation in downtown and central city, and means of transportation solving for metropolitan area are regional line, suburban railway, and intercity railway.

Regional line mainly provides services for metropolitan region (usually suburban areas), and it is generally long radiation and branch-type rail network ingress and egress of the city. Long radioactive lines formed the main traffic flow channel of city; the branch type improves the range of service of regional transit line. The network of regional line is generally rail transit network of medium density, and the average stop spacing is large. Not only it is able to take care of accessibility towns along; but also it can enhance speed to meet the needs of long-distance travel. Therefore, we should realize the regional line and urban line as two systems, and in order to maximize the effectiveness of the entire urban rail transit system, we should let them form closely linked rail system through rational convergence. Otherwise, that will cause difficulties in selecting system format, the functional orientation, and so on.

Influence of Passenger Flow on the Connection of Regional Line and Urban Line

Passenger flow is one of the main factors of regional line and urban line, when convergence is considered. When the connection point of urban line and regional line is not the main passenger hub, transfer passengers will be too large. In this situation, lots of passengers are forced to transfer between regional line and urban line that not only increase the inconvenience of passengers, but also bring pressure on organization of passenger flow at the station.

When the gap of forecasting of line passenger flow intensity is relatively large between the two systems (such as passenger flow intensity of urban line is two times greater than passenger flow intensity of regional line), the length of the train formation may different. Using single point or multi-point transfer connection can simplify the difficulty of coordination of operation management and also that would not cause transfer pressure. In this situation, single point or multi-point transfer connection can be considered [4].

If the passenger flow intensity of regional line is large, it is not suitable to organize off-line train; otherwise, it will bring large impact on urban line. If the passenger flow intensity of urban line is small (peak train interval is greater than three min), organization of off-line train has little effect on urban line, and it can improve the utilization rate of the line through collinear. Under the circumstances, collinear operation can be considered.

Analysis of Connection Mode and Type Between Regional Line and Urban Line

Connecting mode of urban line and the regional line is divided into three modes which including transferring at single point, transferring at multi-points, and running on the same line.

Transferring at Single Point

Connection Mode and Types

Single point transfer refers to the regional line transfer with urban line at terminal site of urban lines or at passenger distribution point of urban and urban fringe, passenger distribution point includes external transport hub and urban public transport hub.

Single Point of Convergence Transfer Problems that Need Attention

1. The regional line and urban line should try to match their capacity each other. Otherwise, it may cause train's unbalanced arrival in the transfer station, that is, in one direction two trains reach at the station, while in the other direction only one train reaches, which will result in passengers' waiting in the station for a period of time, and also it will increase pressure of transfer station.

2. In the scheme of single point transfer, the passenger flow whose destination is downtown is forced to reach the destination after a ride. In the situation, it increases the waiting time of the passengers. Therefore, we have to provide convenient and efficient transfer conditions for passengers at the transfer station. On the one hand, the transfer plan should consider shortening the distance of passengers in the transfer period and shortening the time transfer consumption; On the other hand, the transfer station, the channel, and the escalator should be adapted to a large number of transfer passenger flows. The ideal form is the parallel transfer of the same platform.

Select the Point of Convergence

Because rail transit station should serve the important large-scale passenger distribution point, when the urban rail transit line and the regional line transfer in a single point of transfer, the convergence point generally has two options. You can choose the end of urban rail transit as the transfer point; also, you can choose the line near a large urban area of the passenger terminal distribution point (integrated hub). In the network plans and constructions in the past, peripheral line and urban line are connected often at the end of the line and the edge of the city; it has been proved that it is basically forced transfer; it is not only convenient for passengers, but also causes a heavy passenger flow of the transfer station. Therefore, a single point of transfer should choose transportation hub. In other words, we should transport the regional line passenger to a larger passenger distribution point. Regional line can connect with one or two urban lines (multi-point line), and not exceeding three lines. In this case, the transfer passenger flow is relatively concentrated. Passenger easing and organizational management will not be too complicated.

Transferring at Multi-points

Connection Mode and Types

Transferring at multi-points refers to that regional line introduces to the center of city and crosses with multiple urban lines which can form two and more transfer points. It enlarges the attraction and radiation range of passenger flow. In the situation, regional line forms multiple transfer hubs which is convenient for passengers' travel and transfer, and also it reduces the pressure of transfer passenger flow at the single point. This connection mode has the following two types:

1. Cutting type—Regional line enters the fringe or peripheral area of the city and transfers with urban line through multi-lines and multi-points by cutting with the extension lines of urban rail transit outside of the city center. Then, they will become independent and flexible operation lines of the outer suburban areas.

The passenger section is relatively balanceable in this type, and it is conducive to the dispersion of pressure coming from passenger flow of urban lines.

2. Through type—Regional line enters and passes through the center of the city, and connects with peripheral area or subcenter. At the same time, it can transfer with urban line at multi-lines and multi-points in the city. In the situation, transfer pressure will be big and the line will be very long, so objective speed should be higher. Character of passenger flow is complex, and range of the passenger flow's line section is big, so operational organization and wiring layout are difficult.

Functional Orientation After Regional Line Introduced to the Urban Segment

After the introduction of regional line into urban, according to the different lines in the city area, the function of the line segment is located in the following two types:

1. collecting and dispersing regional passenger flow mainly
   When regional line occupies urban secondary traffic corridors, the functional orientation of the segment extending into downtown will be passengers' transferring, absorbing, and collecting. Therefore, the segment extending into downtown will remain in accordance with the characteristics of the operation of the regional line. In this situation, the average stop spacing is large, and stations are few, so it is necessary to set up stations at the intersection of the main road and the intersection of the urban lines, and the number of stations should not be too much. Thus, the travel speed of the regional line is not affected, and the service level of it can be guaranteed.

2. With the function of urban traffic
   When regional line extends into the center line and occupies main traffic corridors of the city, the function orientation of the regional line is not only collecting and dispersing of the regional passenger flow, but also it must have the function of urban rail transit. In this situation, rapid transit line in metropolitan should give consideration to the function orientation of urban line, the average stop spacing is small, and stations are more. These are contrary to the characteristics of the high speed and long stop spacing of the regional line system. It not only reduces the service level of the regional line, but also brings difficulties to the system selection of the regional line easily.

The Selection of Rout Position and Connection Points After Regional Line Is Introduced to the City Center

1. The selection of rout position after it is introduced to the city center
   Usually, main urban transport corridors have been occupied by the backbone of the urban lines. If regional line occupies the main urban transport corridors after

introduced into urban area, the regional line must provide the function of the urban line, and it is easy to cause that the division of regional line and urban line is not clear. Then, it is easy to form a long line and increase the engineering investment. Therefore, we should try to avoid laying the regional line along the main traffic corridor. If there is a plan that regional line should enter the center of the city, it is necessary to plan corridor for the regional line in urban rail transit network planning or revising.

2. The selection of connection points

The selection of connection point at which regional line enters the urban area should be considered by many factors. The two means that convergence point is at the high-flow section or is in the edge of the urban areas have shortcomings. If the regional line enters into center of the city and connects with the station which has large passenger flow, there may be problems in the ability of the regional line and urban line, or the transport capacity of the urban line may be wasted. If the regional line is connected to the edge of the city, passengers who choose to take it cannot directly reach to the center of the city and have to transfer to another line. In this case, it will reduce the service level of the regional line and will affect the requirement about regional line. Therefore, the choice of the convergence point should be determined according to the actual situation.

Collinear Operation

Connection Modes and Type

The connection mode of collinear operation refers to that in the situation that car, track, line, signal, communication, power supply and other aspects of the regional line are compatible with urban lines, the regional line can share tracks and lines with urban lines. This type mainly includes the following two types.

1. Extended type—because of limited channel resources, regional line will share tracks with the urban lines after its stretching into urban district, or it can be realized as an extension of operating line mode. But the matching of class of passenger flow, vehicles marshaling and the reaching ability of total running time about 1 h must be considered.
2. Mixed operation type—In local district of metropolitan region,in order to sharing corridor resources and avoiding duplication of construction, with the principle of equality and mutual benefit, safety, and reliability, regional line is allowed to share tracks with urban line. In this situation, regional line can be allowed to pass through the stop of urban line without stopping. However, the terminal station and assistant line of the mixed operation section must be clearly defined; also, vehicle of dual power supply system can be applied; and requirements of the transport capacity should be meet according to the unified train operation diagram.

Collinear Adaptation Convergence Condition

Through the above analysis, running with the same line, you must meet the following basic requirements:

1. Infrastructure conditions permit. The bound should meet the requirements of various types of vehicle operation in the area of the same line. The effective length of the station platform and the station wiring need to satisfy different marshaling and different length of train operations; some stations need overtaking lines when organize train overtaking.
2. Power supply, signals, and other standards should be unified or compatible. If line signal, power supply of regional lines and urban lines are not at the same standard, the train cannot run. It is possible to realize running in the same line if we change line technology and make a unified standard, or compatible in specific situation.
3. The segment of running the same line should be fewer passenger flow and smaller rate of traffic flow than the other total amount of passenger lines. Otherwise, it will become a bottleneck for two lines (or more lines) [5]. Domestic and foreign joint operation mode is mainly used in the small traffic density railway [6], suburban railway, or small passenger flow of urban rail transit lines.

Organization of Collinear

Reasonable operating organization program is an effective carrier and basic security of collinear running. Since the city line's running speed is generally different from urban lines, you can organize marshaling or no marshaling program. If not organizing marshaling, there are staggering and limiting-velocity programs [7]. According to the situation, you should choose suitable program of running on the same line to achieve good economic and social benefits. There are mainly three organization programs:

1. Collinear Overtaking
   Collinear overtaking means in collinear segments in which the speed of the city line trains are not limited, city line trains at the station can overtake urban trains. Its advantage is that it can guarantee the city line train traveling speed, thus it can ensure service of the city line's level. But it needs to add more stations, investment of capital construction stations, and other equipment. While at the overtaking stations, it spends longer time on urban train line avoiding the city line which will reduce service level of urban rail transit [8]. In addition, for being influenced by deduction coefficient of mixed operation of high- and medium-speed trains, the carrying capacity of line is reduced in a certain degree.

2. Collinear Staggering

   Collinear Staggering means running urban line during the peak passenger flow and running proper regional rail line during the other period. Its advantage is that it can make the best of the line capacity of sections which can ensure the urban train well run. Thus, it also guarantees the service levels of urban rail transit. However, due to stopping urban line during the peak passenger flow, it will reduce the city line's service level partly.

3. Collinear limit

   Collinear limit means limiting the speed of city line and reducing the level of urban line to making them at the same speed. Its advantage is to ensure not reducing the capacity of the line sections; the service level and quality of the urban line are not affected by the urban line [9]. However, due to the city line and the urban line run at the same speed [10], it will reduce the speed of the city line and influences the passenger's travel time and service level.

Conclusions

Above all, the good connection of regional rail line and urban line is the premise to ensure the healthy development of regional rail line. The connection between the two needs is to consider the basic characteristics of the city, the layered system of urban rail transit, and the influence of the passenger flow. Through comprehensive analysis of connection modes and deeply understanding of the connotation, function positioning, adaptive condition, convergence point and location selection, we can plan and design regional lines correctly.

References

1. Yang, Zhijie. 2013. Analysis on the connection mode of urban rail transit line network. *Comprehensive transportation.* 10: 48–55.
2. Zhe, Zhou., Yang, Cui., and Xinrong, Zhou. 2009. *Study on the connection mode of urban rail transit and urban line.* China Urban Transportation Planning Association 2009, 599–605.
3. Institute of standard quota of ministry of housing and urban-rural development. 2014. *Technical Guideline for Urban Rail Transit System Planning*, 48–55. Beijing: China Building Industry Press.
4. Hu, Xinyu. 2011. *Research on the connection between suburban and urban rail transit in chang sha*, 31–32. China: Southwest Jiao tong University.
5. Gao, Fei., and Lei.li, Yu ping. 2009. Analysis on the connection mode of urban rail transit and regional rail transit. *Railway Transport and Economy* 8: 56–58.
6. Wei, Yun., Qing, Tian., and Teng, Guo. 2013. An improved pedestrian detection algorithm integrating Haar-like Features and HOG Descriptors. *Advances in Mechanical Engineering* 5: 537–543.
7. Gao, Fei. 2010. *Study on the joining mode of urban rail transit and regional rail transit*, 40–41. China: Southwest Jiao tong University.

8. Zhai, Changxu., and Tao, Zhou. 2009. Connection Model of regional railway transportation with Urban rail transit. *Urban Rapid Rail Transit* 22: 36–39.
9. Liu, Haizhou., and Tao, Zhou. 2015. Discussion on connection mode between metropolitan suburban railway and Urban rail network. *Urban Rapid Rail Transit* 28: 53–56.
10. Wei, Yun, Shan Lin, Rui Chu, and Qing Tian. 2016. A method of grading subway stations. *Journal of Transportation Research Procedia* 137: 806–810.

Demand Forecasting-Based Layout Planning of Electric Vehicle Charging Station Locations

**Min Li, Wuhong Wang, Hongfei Mu, Xiaobei Jiang,
Prakash Ranjitkar and Tao Chen**

Abstract The growing use of electric vehicles (EVs) promotes environmental protection and energy conservation. The prerequisite in the use of EVs is that they should be adequately charged. The layout planning of charging station locations is therefore a key point in meeting the charging demands of EVs. This study presents three types of charging demands (i.e., conventional charging, fast charging, and fast battery replacement demand) by forecasting electric vehicle ownership with the use of the Bass model based on traditional vehicle development. This model for locating charging stations is built and optimized on the basis of the forecasted charging demands. The aim is to minimize the layout construction cost for charging station locations and the charging cost for customers. A practical example that applies the model to optimize the layout of charging station locations is presented, and the developed model is validated to work effectively. The model provides a theoretical way to optimize the layout of charging station locations and serves as a basis for layout planners and a reference for other researchers.

M. Li · W. Wang (✉) · H. Mu · X. Jiang (✉)
Department of Transportation Engineering, Beijing Institute of Technology,
5 South Zhongguancun Street, Haidian District, Beijing 100081, China
e-mail: wangwuhong@bit.edu.cn

X. Jiang
e-mail: jiangxiaobei@bit.edu.cn

M. Li
The School of Automobile Qingdao University of Technology,
777 Jialingjiang Road, Huangdao District, Qingdao 266520, China

P. Ranjitkar
Civil and Environmental Engineering, Faculty of Engineering,
University of Auckland, 20 Symonds Street, Auckland 1010, New Zealand

T. Chen
Key Laboratory of Automotive Transportation Safety Techniques of
Ministry of Transport, Chang'an University, Xi'an 710064, China

Introduction

Electric vehicles (EVs) have received much attention because of their pollution-free features. The construction and layout planning of charging stations have become important with the expansion in the use of EVs. The factors that affect charging station locations vary, such as EV ownership and the corresponding charging demands. This study focuses on the optimization of charging station locations on the basis of forecasted charging demands.

Research on charging station locations is relatively mature in foreign countries [1–5]. Namdeo [6] proposed a geography model method in the layout of charging stations according to two constrained conditions, namely people's travel characteristics and intension to use EVs. Xi [7] developed an optimal simulation model for the layout of trickle charge stations to predict the geographical distribution of EV users and simulate the relation between charging speed and charging station layout. Peterson [8] advocated hybrid or pure EVs on the basis of their evaluation of charging costs against those of traditional automobiles. Iversen built a randomized dynamic model that considers EV drivers' characteristics to determine charging station locations. Iversen attributed EV development to the advancement of battery technology and the rise in fuel prices; he also considered accessibility to charging stations and charging cost as key factors that affect people's decision to use EVs [9].

The above studies consider many relative factors and locate charging stations with different targets by using different perspectives. The construction of charging stations is still in its primary stage in China, and it is consistent with the increasing demand to use EVs to minimize construction cost and charging cost for users. This study develops a model to determine the optimal locations for charging stations on the basis of forecasted EV ownership from the perspective of traditional vehicles. The goal is to minimize the total cost of charging stations in the target district. A practical example in Haidian District in Beijing is used to validate the proposed model.

Forecasting Charging Demand

The forecasting of EV ownership is indirectly performed through forecasting of traditional automobile ownership because of data insufficiency on EV market sales.

Forecasting Traditional Automobile Ownership

To analyze the history market data of traditional automobiles according to time series through regression, the future market size can be predicted by the following regression model through trend extrapolation:

$$y_t = a_0 + a_1 t + \cdots + a_k t^k, \tag{2.1}$$

where Y_t is the predicted size of traditional automobiles in millions for the year forecasted. a_0, a_1, ^, and a_k are the regression coefficients that can be calculated by the least squares method. MATLAB can facilitate accurate fitting of these values. t is the time in years.

Forecasting Future EV Ownership

The ownership of EVs, as new industrial, durable products, can be predicted with the Bass model. The discrete form of the Bass model is as follows:

$$\frac{f(t+1)}{1 - F(t)} = p + \frac{q}{m} N(t), \tag{2.2}$$

where $f(t + 1)$ is the ratio of new recipients and all potential recipients at a given point, $F(t)$ is the ratio of the cumulative recipients and all potential recipients at a given point, q is the coefficient for internal influence, which is also called the imitation coefficient, p is the coefficient for external influence, which is also called the innovation coefficient, m is the maximum amount for the potential market, N_t is the cumulative purchases at a given point, and N_t is the purchased amount (not cumulative) at a given point.

From Eq. (2.3),

$$n(t) = mf(t), \tag{2.3}$$

the following equations can be obtained:

$$n(t) = \frac{dN(t)}{dt} = p[m - N(t)] + \frac{q}{m} N(t)[m - N(t)] \tag{2.4}$$

$$f(t+1) = \frac{dF(t)}{dt} = p + (q - p)F(t) - q[F(t)]^2. \tag{2.5}$$

The actual sales of durable products with the Bass model show that the p value can generally be between 0.01 and 0.03. With the value used in the Bass model

when American researchers forecast their EV ownership considered, p is 0.015 and q is 0.35, with a general range of 0.3 and 0.7.

The potential m of the EV market reaches its maximum value when the entire newly purchased traditional automobile is translated to EV sales.

Forecasting the EV Charging Demand

The charging frequency for each EV can be calculated as follows:

$$d = \frac{\overline{LC}}{T \times E_{\text{eff}}}, \tag{2.6}$$

where d is the charging frequency for each EV per day, LC is the average annual mileage of EV in kilometers, T is the average annual travel time in days, and E_{eff} is the available mileage in kilometers, whose value is 120 km.

(1) Forecasting the regular charging demand

The regular charging demand for a charging station per day and the EV amount for the charging station to serve can be calculated as follows:

$$D_{cg} = N_1 \times d_1, \tag{2.7}$$

where D_{cg} is the regular demand of cars per day, N_1 is the amount for regular charging, and d_1 is the charging frequency for each regular charging EV per day.

The average annual mileage is 15,000 km, the average annual time is 260 days, and the available mileage is 120 km. The regular charging frequency d_1 is 0.48 per day.

(2) Forecasting the fast charging demand

The fast charging demand for a charging station per day and the EV amount for the fast charging station to serve can be calculated as follows:

$$D_{kc} = N_2 \times d_2 + N_1 \times f, \tag{2.8}$$

where D_{kc} is the fast charging demand for each station per day in vehicles per day, N_2 is the fast charging amount of an electric taxi, d_2 is the charging frequency for each taxi to fast-charge per day, N_1 is the fast charging amount of a private EV, and f is the charging frequency of private EVs to fast-charge per day.

The average annual mileage of an electric taxi is 1 million kilometers, its average annual travel time is 350 days, and its available mileage is 120 km. Equation (2.8) shows that the charging frequency of an electric taxi is 2.38 per day, whereas that of an electric private EV is 0.1 per day.

(3) Forecasting fast battery replacement demand

The fast battery replacement demand of a charging station and the number of electric bus for a fast battery replacement station to serve can be forecasted as follows:

$$D_{kh} = N_3 \times d_3. \tag{2.9}$$

In the formula, D_{kh} is the fast battery replacement demand per day, N_3 is the number of electric bus to fast-replace battery, and d_3 is the replacement frequency for the electric bus to fast-replace battery per day.

The average annual mileage of an electric bus is 70 thousand kilometers, its average annual travel time is 350 days, and its available mileage is 250 km. Equation (2.9) shows that the fast battery replacement frequency is 0.8 per day.

Building the Model to Layout the Charging Stations

Locating the Charging Station in the Model Follows the Principles Below

(1) The target district in the layout of the charging station is manually divided into small regions, and the charging demand of each region is located in the centroid of the target region.
(2) The customers of each region can only select the nearest charging station to charge.
(3) The initial selection of the candidate location is consistent with the charging demand distribution.
(4) The charging cost is linearly correlated with the distance to the charging station.

Model to Build

The building model attempts to minimize the annual cost of the charging station as follows:

$$\min T = F_1 + F_2 + F_3, \tag{3.1}$$

where T is the annual cost of the charging station, F_1 is the annual construction cost, F_2 is the annual running cost, and F_3 is the charging cost for the customer.

The annual cost of the charging station in the model includes the construction cost, the running cost, and the customer's charging cost. The customer's charging cost is the generated cost in the charging progress, which is composed of the consumed electric power and the time forward to the charging station.

(1) Annual construction cost

The annual construction cost includes the cost to build the infrastructure of the charging station and the charging pile. It can be calculated as follows:

$$F_1 = p \times \frac{a(1+a)^n}{(1+a)^n - 1} \times JC, \tag{3.2}$$

where p is the cost of the proposed charging station, a is the return rate of investment or the discount rate, n is the estimated investment recovery period by year, and JC is the cost of infrastructure per station.

The model considers the main investments to build a charging station. The coefficient of annual construction cost translates the total construction investment of a charging station into the annual average cost.

(2) Annual running cost of a charging station

The annual running cost is forecasted by the rate of running cost in the total investment, which can be valued according to the rate of other similar public infrastructure. Therefore, the annual average running cost is calculated as follows:

$$F_2 = \mu F_1 \tag{3.3}$$

where u is the rate coefficient.

(3) Cost for customers to charge

As a part of the charging cost for customers, the consumed electric power and time forward to charge are mainly related to the distance between the station and the site, which increases the charging demand and the traffic congestion along the distance. The basic principles of the model show that the charging cost for customers is proportional to the distance as follows:

$$F_3 = wm \sum_i \sum_j D_i Q_{ij} Z_{ij} d_{ij}, \tag{3.4}$$

where w is the charging cost for the customer per kilometer, m is the buckling coefficient of the path to the charging station, D_i is the EV amount to charge, Q_{ij} is the choice matrix, Z_{ij} is the coefficient matrix for traffic congestion, and d_{ij} is the matrix of distance for the demand sites to the proposed charging stations.

The charging cost for customers is the consumed cost per kilometer for customers to locate a charging station. It is a Euclidian distance from the demand site to the proposed charging station. The actual road network is similar to a chessboard,

so the buckling coefficient is introduced for translation. The choice matrix is a 0–1 matrix, where the element is valued at 1 when the corresponding demand site reaches the corresponding charging station and 0 otherwise. The coefficient matrix of traffic congestion is determined by the actual traffic of the planning district. A value of 1 means smooth traffic, 2 means slow traffic, and 3 means jam traffic for this coefficient. The traffic state of the road is the average value of the states all day.

(4) Annual cost of a charging station

The above analysis shows that the mathematical model for the annual minimum cost of the charging station is as follows:

$$\min T = F_1 + F_2 + F_3$$

$$= (1 + \mu)p \frac{a(1+a)^n}{(1+a)^n - 1} \times JC \tag{3.5}$$

$$= wm \sum_i \sum_j D_i Q_{ij} Z_{ij} d_{ij}$$

$$\sum_j Q_{ij} = 1, \quad \forall i \in I \tag{3.6}$$

$$Q_{ij} \leq X_j, \quad \forall i \in I, \quad \forall j \in J \tag{3.7}$$

$$\sum_j X_j = p \tag{3.8}$$

$$d_{ij} \leq R_j, \quad \forall i \in I, \quad \forall j \in J \tag{3.9}$$

$$\sum_i D_i Q_{ij} \leq C X_j, \quad \forall j \in J \tag{3.10}$$

$$X_j \in \{0, 1\}, \quad \forall j \in J \tag{3.11}$$

$$Q_{ij} \in \{0, 1\}, \quad \forall i \in I, \quad \forall j \in J \tag{3.12}$$

where C is the service level of the charging station, I is the set of demand sites, and J is the set of proposed charging stations.

All of these constraints can be explained as follows. Equation (3.6) shows that a consumer in every district can only select the same charging station. Equation (3.7) indicates that the consumer can only charge when one station is selected. Equation (3.8) shows the total amount of proposed charging stations. Equation (3.9) indicates that the distance from the demand site to the station should be shorter than the serving radius. Equation (3.10) shows that the EV amount to charge should be smaller than the service capacity of the charging station. Equation (3.11) is a 0–1 vector of candidate stations. Equation (3.28) is a matrix

where elements are valued as 1 or 0, which indicate whether the EV goes to the corresponding station.

Practical Example

The planning layout of Haidian District, Beijing City, is selected as a practical example. The demand district is divided into administrative areas of streets, villages, and towns for convenience. These divisions are numbered from 1 to 24.

Forecasting the Number of EVs

Because of the lack of immediate related data on Haidian District to forecast the charging demand, the per capita EV ownership in Beijing and the population of the target district should be predicted first.

(1) Forecasting the EV ownership in Beijing

The available history data of traditional motor vehicles are listed in Table 1. The future ownership of traditional motor vehicles is predicted by fitting of curves in Table 2, and the largest potential market size of EVs in every future year is fixed by the growth of traditional motor vehicles. In this prediction, the significant influence of government policy on private vehicles from 2011 has been temporarily ignored for theoretical discussion. Taxi ownership is predicted to grow by 200 units annually because of the strict limitation set by government policy.

Private EV ownership by the year 2020 can be forecasted according to Eq. (2.3) of the Bass model from the EV ownership in 2013. The consistent policy support shows that the number of electric taxis and buses will account for 30% of the total. All predicted data are shown in Table 3.

Table 1 History data of traditional motor vehicles

Year	Private car (thousand)	Taxi	Bus
2003	1056	62,283	16,170
2004	1298	51,561	18,612
2005	1540	66,000	18,801
2006	1810	66,646	19,552
2007	2121	66,646	19,395
2008	2483	66,646	21,507
2009	3003	66,646	21,716
2010	3744	66,646	21,548
2011	3897	66,646	21,628
2012	4075	66,646	22,146

Table 2 Traditional motor vehicles are predicted by fitting of curves

Year	Predicted private car	Predicted taxi	Predicted bus
2013	5851	67,646	22,237
2014	6729	67,846	22,403
2015	7680	68,046	22,590
2016	8704	68,246	22,823
2017	9803	68,446	23,123
2018	10,974	68,646	23,516
2019	12,220	68,846	24,024
2020	13,539	69,046	24,671

Table 3 Consistent policy support

Year	$f(t)$ (%)	$F(t)$ (%)	Electric private car
2013	–	0.421	3388
2014	1.641	1.058	17,792
2015	1.850	1.344	35,390
2016	1.944	1.512	55,296
2017	1.999	1.624	77,260
2018	2.035	1.706	101,089
2019	2.061	1.767	126,771
2020	2.081	1.816	154,221

Table 4 Per capita EV ownership in Beijing

Electric private cars	Electric taxis	Electric buses
49.9 for every 10,000 persons	6.7 for every 10,000 persons	2.4 for every 10,000 persons

Table 5 Ownership of the target district

Electric private cars	Electric taxis	Electric buses
22,435	3012	1079

(2) Forecasting the EV ownership in Haidian District

The population of Beijing in 2020 is predicted to reach about 30 million, as identified from fitting the population from 2001 to 2012 in the 2013 Beijing statistical yearbook.

The above prediction shows that the per capita EV ownership in Beijing is fixed, as shown in Table 4.

The population of Haidian District in 2020 can be fitted from historical data at 4.496 million. The ownership of the target district is shown in Table 5.

Private EV and taxi ownership in these regions are shown in Table 6 according to the population ratio of every administrative region in the target district.

Table 6 Distribution of electrical buses focuses on the public transport

Numbered area	Electric private cars	Electric taxis	Numbered area	Electric private cars	Electric taxis
D1	1301	175	D2	960	129
D3	370	50	D4	810	109
D5	1030	138	D6	893	120
D7	1286	173	D8	772	104
D9	1043	140	D10	1203	161
D11	1528	205	D12	1097	147
D13	1209	162	D14	1126	151
D15	211	28	D16	287	39
D17	393	53	D18	1837	247
D19	976	131	D20	379	51
D21	801	108	D22	772	104
D23	1059	142	D24	1090	146

Table 7 Three types of charging demand per day

Regular charging demand per day, D_{cg}	Fast charging demand per day, D_{kc}	Fast battery replacement demand per day, D_{cg}
10,769	9412	1802

The distribution of electrical buses focuses on the public transport hub because it is uninfluenced by population distribution.

Calculating the Charging Demand

Equations (2.7), (2.8), and (2.9) show that the three types of charging demand per day in the target district can be calculated as follows (Table 7).

The fast charging demand is the main target to plan the charging station for two reasons: Customers regularly charge in their garage, and the fast battery replacement demand can be satisfied in the focused transport hub for electric buses. Therefore, the fast charging demand of each area D_i can be calculated as follows:

$$D_i = 2.38D_{cz} + 0.1D_{sj}, \tag{4.1}$$

where D_{cz} is the ownership of electric taxis in the area, and D_{sj} is the ownership of private electric cars.

Table 8 shows the results of the fast charging demand in every target area.

Table 8 Results of the fast charging demand in every target area

Numbered area	Fast charging demand	Numbered area	Fast charging demand
D1	547	D2	506
D3	403	D4	472
D5	156	D6	88
D7	340	D8	122
D9	431	D10	165
D11	375	D12	772
D13	540	D14	409
D15	325	D16	159
D17	438	D18	337
D19	503	D20	325
D21	641	D22	444
D23	460	D24	456

Fig. 1 Fourteen stations from the proposed 48 stations

Locating the Charging Station

The final result (Fig. 1) shows 14 stations from the proposed 48 stations based on the proposed model and with the use of Lingo software.

Conclusion

Charging stations should be constructed to develop EVs. The construction of charging stations speeds up EV promotion. A planner can develop the layout of charging stations on the basis of charging demands in the future. This study presents three main types of charging demands that are predicted by forecasting EV ownership. An optimal model is developed on the basis of the forecasted charging demand to minimize the total cost of the charging stations and optimize the locations of the charging stations in the target district. Haidian District is taken as a practical example, and it indicates that the model can be effectively applied and can determine the location plan of charging stations with optimization. The proposed model can provide a practical way for planners to optimize the locations of charging stations. Some factors excluded in forecasting EV ownership can be considered in future research to improve model precision.

References

1. Joana, Cavadas., Goncalo, Correia., Joao, Gouveia. 2014. Electric vehicles charging network planning. Computer-based modelling and optimization in transportation. *Advances in Intelligent Systems and Computing* 262: 85–100.
2. Jaeyoung, Jung., Joseph, Y. J. Chow., R. Jayakrishnan., and Ji Yong, Park. 2014. Stochastic dynamic itinerary interception refueling location problem with queue delay for electric taxi charging stations. *Transportation Research Part C: Emerging Technologies* 40: 123–142.
3. Aviral, Shukla., Joseph, Pekny., Venkat, Venkatasubramanian. 2011. An optimization framework for cost effective design of refueling station infrastructure for alternative fuel vehicles. *Computers & Chemical Engineering* 35 (8) (10 August 2011): 1431–1438.
4. Phonrattanasak, P; Leeprechanon, N. 2014. Multiobjective ant colony optimization for fast charging stations planning in residential area. In Innovative Smart Grid Technologies-Asia (ISGT Asia), 2014 IEEE.
5. Zheng, Yu-Jun., Chen, Sheng-Yong. 2013. Cooperative particle swam optimization for multiobjective transportation planning. *Applied Intelligence* 39: 202–216.
6. Namdeo, A., A. Tiwary., R. Dziurla. 2013. Spatial planning of public charging points using multi-dimensional analysis of early adopters of electric vehicles for a city region. In *Technological Forecasting and Social Change*. Available online 8 October 2013, ISSN 0040-1625.
7. Xi, X., R. Sioshansi, and V. Marano. 2013. Simulation–optimization model for location of a public electric vehicle charging infrastructure. *Transportation Research Part D: Transport and Environment* 22: 60–69.
8. Scott, B., Peterson., and Jeremy J. Michalek. 2013. Cost-effectiveness of plug-in hybrid electric vehicle battery capacity and charging infrastructure investment for reducing US gasoline consumption. *Energy Policy* 52: 429–438.

9. Sanzhong, Bai., Du, Yu., and Lukic, S. 2010. Optimum design of an EV/PHEV charging station with DC bus and storage system. In *Energy Conversion Congress and Exposition (ECCE)*. IEEE.

10. Nakul, Sathaye, and Kelley Scott. 2013. An approach for the optimal planning of electric vehicle infrastructure for highway corridors. *Transportation Research Part E-Logistics and Transportation Review* 59: 15–33.

11. He, F., D. Wu, Y. Yin, et al. 2013. Optimal deployment of public charging stations for plug-in hybrid electric vehicles. *Transportation Research Part B: Methodological* 47: 87–101.

12. Obama, B. 2008. Energy Speech Fact Sheet. http://www.barackobama.com/pdf/factsheet_energy_speech_080308.pdf

13. Hisatomo, Hanabusa., Ryota, Horiguchi. 2011. A study of the analytical method for the location planning of charging stations for electric vehicles. *Knowledge-Based and Intelligent Information and Engineering Systems Lecture Notes in Computer Science* 6883: 596–605.

14. Iversen, E. B., Morales, J. M., and Madsen, H. 2013. Optimal charging of an electric vehicle using a markov decision process. 1310. 6926.

15. Becker, Y. A., Tenderich, B. 2009. Electric vehicles in the United States: a new model with forecasts to 2030. In *Center for Entrepreneurship & Technology (CET)*, 2009, 2009. 1.v.2.0.

Geometric Safety Design of Freeway off-Ramp-Street Terminal Based on Traffic Flow Characteristic Analysis

Hai-juan Zhao and Hai-ping Zhao

Abstract Off-ramps are the sites of far more crashes and lower capacity than other segments of freeways, thus appropriate design method is of great significance to improve the safety and capacity performance. This paper described a study on the geometric safety design method of freeway off-ramp-street terminal based on traffic flow characteristic analysis. A field survey was conducted in 17 cities in order to investigate the difference of traffic flow characteristics between the terminal and common intersections. Experimental results indicated that continuous traffic flow and high vehicle speed were the most significant features of the terminal. Moreover, several key points of geometric safety design of the terminal were proposed according to the traffic flow characteristics and driving behavior. Finally, a freeway off-ramp-street terminal in Shanxi Province was selected to promote the application of the geometric safety design method. Case study results have shown the promising property of proposed approach to assist transportation professionals in properly designing the terminal.

Keywords Off ramp-street terminal · Geometric safety design · Traffic flow characteristics · Off-ramp · Freeway

Off-ramps are the sites of far more crashes and lower capacity than other segments of freeways, thus appropriate design method is of great significance to improve the safety and capacity performance. This paper described a study on the geometric safety design method of freeway off-ramp-street terminal based on traffic flow characteristic analysis. A field survey was conducted in 17 cities in order to investigate the difference of traffic flow characteristics between the terminal and common intersections. Experimental results indicated that continuous traffic flow

H. Zhao (✉)
Wuhan Planning and Design Institute, Wuhan 430014, Hubei, China
e-mail: zhaohaijun@wpdi.cn

H. Zhao
Land Resources and Planning Bureau of Wuhan East Lake High-tech Zone,
Wuhan 430079, Hubei, China

and high vehicle speed were the most significant features of the terminal. Moreover, several key points of geometric safety design of the terminal were proposed according to the traffic flow characteristics and driving behavior. Finally, a freeway off-ramp-street terminal in Shanxi Province was selected to promote the application of the geometric safety design method. Case study results have shown the promising property of proposed approach to assist transportation professionals in properly designing the terminal.

Introduction

Off-ramp-street terminal is the key transfer station for the traffic flow of freeway and the surface street. Due to the disturbance of deceleration and lane-changing behavior between going straight vehicles and right-turning vehicles, the capacity of freeway off-ramp significantly decreases, and the traffic flow characteristics become extremely complex [1]. In addition, the off-ramp queens frequently spill over onto the freeway and then cause traffic bottleneck with a diminished capacity [2]. It is a tough job to deal with traffic congestion originated from off-ramps because there is no direct way to control the freeway exit flow [3]. Therefore, drivers are likely to have an elevated crash risk when driving on off-ramps compared with common sections of freeways. It has been reported that a substantial proportion of total freeway crashes occur on ramps and in the vicinity of ramps. Take California, for example, ramp-related crashes accounted for approximately 15% of total freeway crashes during 1992–1994 [4]. According to the US Fatality Analysis Reporting System (FARS) and General Estimates System (GES), an estimated 82,609 police-reported crashes were interchange-related on interstates in 2001, of which nearly 83% occurred specifically on entrance or exit ramps [5].

Freeway ramp has been a hot topic of considerable interest to researchers. Chen and Zhou [6] described a study evaluating the safety performance of left-side off-ramps at diverge areas; they established a crash prediction model to investigate the factors that contribute to the crashes, and experimental results betrayed that the left-side off-ramp always led to higher average crash counts, crash rate, and percentage of severe crashes. Günther and Coeymans [7] proposed a surface streets coordinated method to mitigate freeway off-ramp congestion, and the competing vehicle flow was determined to be detoured to underutilized roads in the local network in order to avoid the formation of queues on freeway off-ramps. The control strategy significantly improves capacity and reduces total delays, however, succeeds at the cost of surface streets users. Chen and Liu [8] investigated the influences of different number and arrangement of lanes on freeway off-ramps at diverge areas, and it was found that a lane-balanced exit ramp resulted in a 68.33% reduction of crash counts by using the crash prediction model. St-Aubin and Miranda-Moreno [9] presented a method for surrogate safety analysis at protected freeway ramps; they found empirically that rear-end interactions were more predominant than lane-changing interactions on freeway ramps. Jin and Wang [10]

developed a variable timing control strategy of surface street to reduce collisions between freeway and surface street vehicles, and simulation results indicated that the approach significantly outperforms common arrangement especially during peak traffic flows. Xie and Ma [11] conducted an analysis of single-lane right exit ramp on freeways; the study came to a conclusion that the parallel type was safer and of a higher capacity compared with the taper type.

As is readily apparent in the above brief review, a wide range of methods and measures have been used to study freeway ramps. Most efforts to investigate freeway ramps have focused on how to improve capacity and reduce crashes through traffic control strategy, but the traffic flow characteristics of off-ramp-street terminal have been the subject of relatively little recent research. Besides, these terminals are always considered as common intersections and use the same design method and management. In this study, we proposed a geometric safety design method based on traffic flow characteristic analysis, and the approach was applied to a sample of off-ramp-street terminal in Shanxi Province, China.

Methodology

Traffic Flow Characteristics of Terminal

In order to investigate the difference of traffic flow characteristics between the terminal and common intersections, we conducted a field survey of off-ramp-street terminals in 17 cities, including Nanjing, Yangzhou, Xuzhou, Suzhou, and so on. The video camera was used to collect data of traffic volume and vehicle headway, and the measurement of vehicle speed was greatly simplified by using the radar velocimeter. The sufficiently large video data was analyzed on the basis of manual observation and mathematical statistics, and the traffic flow characteristics were summarized from the perspective of extracting difference between the terminal and common intersections.

(1) Continuous Traffic Flow of Terminal

Vehicles of suburban freeways and mainline toll-controlled intercity freeways always arrive at the terminal randomly and discretely. Furthermore, surface street vehicles are always of priority compared with unsignalized terminal vehicles, and the survey results indicate no apparent acceleration or deceleration when crossing the terminal, namely the traffic flow of off-ramp-street terminal can be considered as continuous flow. On the contrary, the traffic flow of common intersections may be discontinuous due to traffic management, vehicle interactions, and many other factors.

(2) High Vehicle Speed on Terminal

Drivers are likely to keep a high speed when they drive through the off-ramps because of driving misconception, and this phenomenon becomes much more prominent referring to overpass freeway interchanges. Vehicle speed was
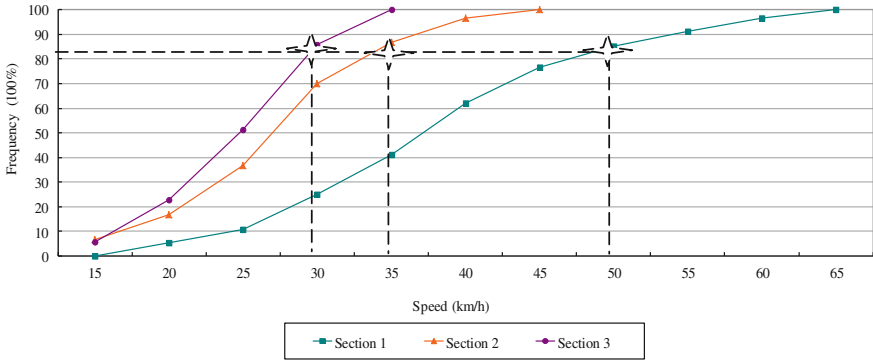
Fig. 1 Speed distribution of interchange off-ramp vehicles on different points
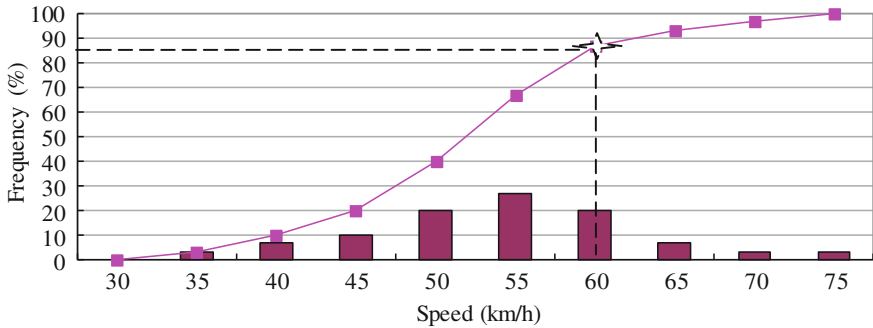

Fig. 2 Speed of crossroad vehicles around road

investigated on the off-ramp with posted speed limit of 40 km/h, and the speed of 3 sections was measured by using the radar velocimeter. The 85th percentile speed of Section "Introduction" (50 m away from the stop line), Section "Methodology" (30 m away from the stop line), and Section "Case Study" (stop line) was 50, 35, and 30 km/h, respectively (Fig. 1). Furthermore, going straight vehicles of surface streets arrived at the terminal at a speed of nearly 60 km/h, and there was no apparent deceleration (Fig. 2).

(3) Small Gap Acceptance on Terminal

Experimental results revealed that the gap acceptance on terminal was smaller than common intersections because of high speed and traffic volume. Off-ramp vehicles constantly force an entrance into surface streets and then cause traffic congestion, especially when left-turning trucks break into the streets. The terminal results in more reckless driving, namely drivers are likely to accept smaller gaps that should have been rejected in common intersections (Table 1). An inaccurate assessment of gap acceptance can probably increase the driving risk and even lead to a traffic accident [12].

Table 1 The gap acceptance for off-ramp vehicles turning into surface streets

Type of vehicles	Passenger cars	Middle size vehicles	Oversize vehicles	Articulated vehicles
Right-turning t_{cr}(s)	2.5	3.0	4.0	5.5
Left-turning t_{cl}(s)	3.0	4.0	5.0	6.0

(4) Considerable Proportion of Trucks on Terminal

Trucks account for a large proportion of freeway traffic volume and have intense interference effect on surrounding vehicles. It was claimed that a 30% increase of heavy vehicles resulted in a 5% increase of accident probability [13]. Therefore, trucks should be taken into account in channelization design of the terminal.

(5) Significant Impact of Road Grade

The road grade has a significant impact on vehicle speed, thus a full consideration should be given to road grade in geometric design of off-ramp-street terminal. Apparent acceleration or deceleration was found on terminals when the off-ramp was connected downgrade with the surface streets, as opposed to the upgrade access management.

Geometric Safety Design of Terminal

Site Selection of Terminal

(1) Selection of Surrounding Land Use Patterns

Off-ramps are essentially connected with many traffic zones rather than the surface streets, thus the traffic flow characteristics of terminal are largely determined by surrounding land use patterns: life-oriented or traffic-oriented and passenger-oriented or freight-oriented. Land use patterns were divided into four types from the perspective of traffic generation, including commercial land, entertainment land, residential land, and industrial land.

Commercial land, entertainment land, and residential land are traffic attractive points with high development density; most roads around these lands are applied to meet the needs of economic communication, administrative management, cultural education, and medical treatment. However, the traffic flow generated by industrial land often serves production and daily life of a large area. Industrial land has greater demand for convenient arterial highways with less access roads and intersections than other lands. Therefore, freeway off-ramps should be as far as possible away from commercial land, entertainment land, and residential land, while the industrial land may contribute to traffic gather and distribution.

(2) Selection of Access Roads

The access road of off-ramps should not be too far away from important traffic source with the purpose of convenient gather and distribution of traffic flow. In order to disperse the traffic flow of off-ramps rapidly, access roads are required to have spare capacity and the main road is recommended to be chosen. In addition, there should be less branch roads and buildings around the access road. However, shopping malls and central plaza are often found to be situated near the off-ramps in China [14]. Closure and combination management are proved to be effective in dealing with too close openings, and the new opening ought to be set in main intersection which is far away from the access road.

Further analysis of feasibility is necessary after preliminary determination of the access road, and the impact on surrounding traffic network is fully considered based on microcosmic analysis and macroscopic evaluation method. Microcosmic index includes intersection saturability, spilled over queen length, and average delay, while macroscopic evaluation mainly focus on traffic uniformity, service level, and flow distribution from the perspective of whole network. This paper only provides a thought on this topic, because extensive research has been conducted in the past few decades.

(3) Selection of Upstream and Downstream Intersections

Upstream and downstream intersections have a significant influence on safety and capacity performance of the terminal. The distance between upstream and downstream intersections should first meet the requirements of common intersections. Too shorter spacing may result in higher risk of driving, and we recommend the distance to be at least 500 m on suburban freeways and at least 1000 m on intercity freeways. It is not reasonable to set too many access roads between the two intersections, and generally speaking, the number of left-unlimited road should be less than two. Furthermore, upstream and downstream intersections are required to have higher capacity in order to disperse the traffic flow of off-ramps.

Key Points of Geometric Safety Design

According to the traffic flow characteristics and common design errors of off-ramp-street terminal, key points of channelization design method were presented to improve its safety performance.

(1) Based on Traffic Volume

Separated island, channelized island, and some other measures are used to regulate driving route and organize traffic flow according to the actual demands. Terminal channelization should be based on the turning traffic volume of freeway off-ramps and take full consideration to vehicle speed, road grade, crossing angle, interchange form, control method, and so on.

(2) Focus on Left-turning Vehicles

The organization of left-turning vehicles is of great importance in terminal channelization, it is reasonable to make it orthometric crossing as far as possible, and left-turning lane should be provided when necessary. Besides, the Y-shaped terminal should be avoided, because of the heterotropic crossing of left-turning vehicles.

(3) Ensure Adequate Sight Distance

Channelization design is aimed to guide drivers to choose appropriate driving route on terminals; hence, channelized facilities and green landscape ought not to obstruct the view of drivers, and an adequate sight distance is needed, especially on underpass freeways.

(4) Take Concise Channelization Measures

Channelized islands and traffic markings must be entirely compliant with related design specification so as to remind drivers of proper route. It is not easy for drivers to correctly understand, especially when they are driving at a high speed on terminals, so channelization measures should be concise and comprehensive with no misunderstandings. For the purpose of safety driving, channelization should be driver-oriented rather than designer-oriented.

(5) Pay Attention to Treatment of Different Terminals

Interchange terminals usually appear in pairs, for instance, terminals and intersections often concentrate when rhombus or cloverleaf interchanges connected with surface streets. The form and necessary lane number of channelization should be fully considered because of the short distance between the terminals and intersections.

(6) Focus on Detail Design

The form and size of separated island, channelized island, and left-turning lane are determined by traffic operation and safety demands. For example, it is easy to be ignored that offset and sleek treatment are needed on the incident flow boundary of channelized island.

(7) Meet the Requirements of Oversize Vehicles

The traffic characteristics of oversize vehicles should be taken into consideration if the oversize vehicles account for a large proportion of turning flow. In addition, steep downgrade, small turning radius, and reverse superelevation should be avoided in terminal design.

(8) Deal with Heterotropic Terminals

Orthometric terminals provide left-turning vehicles the shortest route to pass the lanes, and unnatural turning curves are eliminated at the same time. Separated island and channelized island may be necessary if the crossing angle of terminals is less than 70 degree.

Case Study

Current Situation of the Terminal

The terminal is situated on Fen-jun freeway off-ramp, connected with 340 provincial highway (S340) and 307 national highway (G307) in Shanxi Province, China (Fig. 3). Fen-jun freeway off-ramp is a long downgrade joined with the road S340 in the northeast, and it is 8 m wide with no lane division. Road S340 is the first grade highway with six lanes and a width of 25 m, while road G307 is the second grade highway with two lanes and a width of 15 m. Clear distance of the terminal is 120 m from east to west and 60 m from south to north. Unsignalized management and a serious lack of channelization result in disordered driving and even cause vehicle converse at times. Besides, road capacity is severely decreased by pedestrian violations and intensive interference from the adjacent logistics company, gas station, and some other organizations.

Geometric Safety Design of the Terminal

According to the current situation of the terminal, geometric safety design method was used to improve the safety and capacity performance (Figs. 4 and 5).

(1) The road alignment of G307 was optimized and the crossing angle of S340 was enlarged in order to make it a T-shaped intersection.
(2) The terminal was redesigned by increasing the kerb turning radius and shrinking the physical area, and the four-lane S340 was widened to be six lanes so as to increase capacity and reduce traffic congestion. Left-turning lane, right-turning lane, and traffic markings were used to define right of way on the
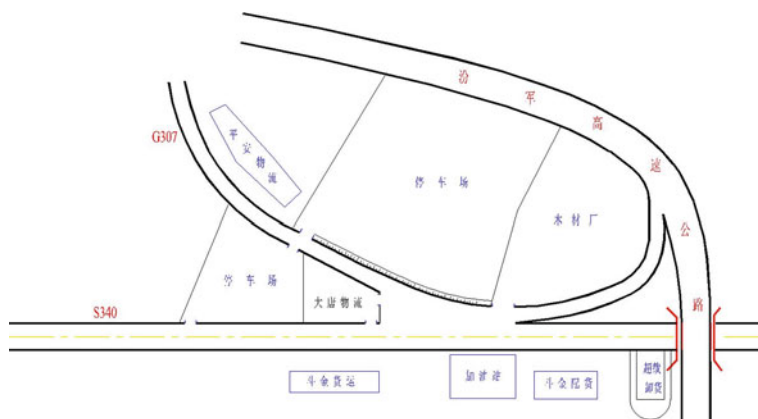
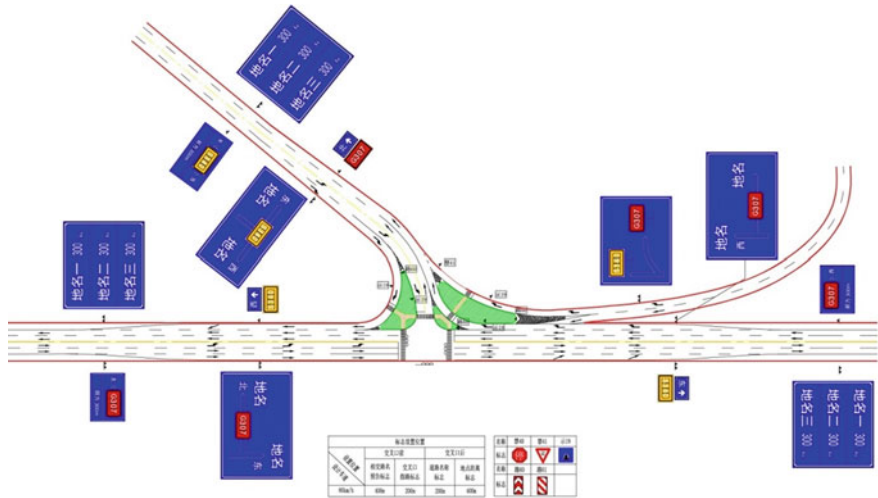**Fig. 3** Current situation of the terminal

Fig. 4 Geometric safety design of the terminal

terminal. Channelized islands were implemented on both on-ramp and off-ramp of 307 national highway. Furthermore, guide sign, pedestrian crossing sign, and two-way traffic sign were provided to organize the traffic flow in order.

(3) Closure management was used to deal with the openings over against the terminal, such as Da Tang company and the timber mill. Other adjacent openings were combined to be a new one so as to weaken the interference.

(4) The terminal was redesigned to be a signalized intersection with both vehicle signal and pedestrian signal. The signal control plan proceeds as follows: going straight vehicles share one phase with left-turning vehicles in the west of S340 (Phase 1); left-turning vehicles from off-ramp share one phase with going straight and right-turning vehicles in the east of S340 (Phase 2); left-turning vehicles share one phase with right-turning vehicles on G307 (Phase 3).

(5) Pedestrian crossing was implemented to protect vulnerable road users, and colorized ground tile was advocated on channelized islands.

Evaluation of the Improvement

Channelized islands, lane division, T-shaped intersection, and road alignment optimization succeeded in reducing traffic conflicts from the aspect of space, while the signal control plan achieved this goal from the aspect of time (Table 2). Pedestrian safety was greatly enhanced due to pedestrian crossing and signal timing. Furthermore, the hard shoulder of S340 was fully used in order to widen traffic lanes, thus capacity of the terminal was improved to a certain degree.

Before improvement of provincial highway S340   After improvement of provincial highway S340

Before improvement of national highway G307   After improvement of national highwayG307

Fig. 5 Before and after comparison on improvement of the terminal

Table 2 Before and after comparison ontraffic conflicts of the terminal

Type of conflicts	The number of traffic conflicts			
	Before improvement	After improvement		
		Phase 1	Phase 2	Phase 3
Crossing	11	0	0	0
Merging	6	0	0	0
Diverging	5	0	0	0

Conclusion

Appropriate geometric safety design method of off-ramp-street terminal is significant for the improvement of safety and capacity performance. This paper described a study investigating the traffic flow characteristics of the terminal, and experimental results showed a significant difference between the terminal and common

intersections. Continuous traffic flow and high vehicle speed were found to be the most significant features of the terminal. Proposed approach was applied to the improvement of off-ramp-street terminal on Fen-jun freeway, and it was found that traffic conflicts (crossing, merging and diverging conflicts) were eliminated by using the geometric safety design method, namely safety performance of the terminal was successfully improved.

Despite the success of presenting an effective design method of the terminal, there were several limitations in the study and some of the very basic questions need further research. On the one hand, only one terminal was investigated due to the limitation of time, manpower, and material resources. On the other hand, much more remains to be carried out in the future study regarding the evaluation method of the improving effect. However, at least, it sheds some light on the topic. Subsequent research will focus on different forms of terminals and benefit evaluation method of the improvement.

References

1. Gong, J. G., X. C. Guo, and S. Dai, et al. 2013. Research on the Operational Characteristics of Diverging Area on Expressway Off-ramp Based on Different Restricted Strategy of Lane. *Procedia- Social and Behavioral Sciences* 96:1156–1164.
2. Cassidy, M. J., S. B. Anani, and J. M. Haigwood. 2002. Study of Freeway Traffic Nearan Off-Ramp. *Transportation Research Part A: Policy and Practice* 36 (6):563–572.
3. Spiliopoulou, A., M. Kontorinaki, and M. Papageorgiou, et al. 2014. Macroscopic Traffic Flow Model Validation at Congested Freeway Off-Ramp Areas. *Transportation Research Part C: Emerging Technologies* 41:18–29.
4. Lee, C., and M. AbdelAty. 2009. Analysis of Crashes on Freeway Ramps by Location of Crash and Presence of Advisory Speed Signs. *Journal of Transportation Safety & Security* 1 (2):121–134.
5. McCartt, A. T., V. S. Northrup, R. A. Retting. 2004. Types and Characteristics of Ramp-Related Motor Vehicle Crashes on Urban Interstate Roadways in Northern Virginia. *Journal of Safety Research* 35 (1):107–114.
6. Chen, H. Y., H. G. Zhou, and J. G. Zhao, et al. 2011. Safety Performance Evaluation of Left-Side Off-Ramps at Freeway Diverge Areas. *Accident Analysis & Prevention* 43 (3): 605–612.
7. Günther, G., J. E. Coeymans, and J. C. Muñoz, et al. 2012. Mitigating Freeway Off-Ramp Congestion: A Surface Streets Coordinated Approach. *Transportation Research Part C: Emerging Technologies* 20 (1):112–125.
8. Chen, H. Y., P. Liu, and J. J. Lu, et al. Evaluating the Safety Impacts of the Number and Arrangement of Lanes on Freeway Exit Ramps. *Accident Analysis & Prevention* 41 (3): 543–551.
9. St-Aubin, P., L. Miranda-Moreno, and N. Saunier. 2013. An Automated Surrogate Safety Analysis at Protected Highway Ramps Using Cross-Sectional and Before-After Video Data. *Transportation Research Part C: Emerging Technologies* 36: 284–295.
10. Jin, S., D.H. Wang, and L. Wang. 2009. Traffic Control Strategy for a Surface Street on an Expressway-Arterial Corridor. *Tsinghua Science & Technology* 14 (6): 776–781.

11. Xie, J. P., Y. F. Ma, and L. Yuan, et al. 2016. Safety and Capacity Performances of Single-Lane Right Exit Ramp on Freeway: A Case Study in Jiangsu Province, China. *Procedia Engineering* 137:563–570.
12. Hunt, M., D. N. Harper, and C. Lie. 2011. Mind the Gap: Training Road Users to Use Speed and Distance When Making Gap-Acceptance Decisions. *Accident Analysis & Prevention* 43 (6):2015–2023.
13. Moridpour, S., E. Mazloumi, and M. Mesbah. 2015. Impact of Heavy Vehicles On Surrounding Traffic Characteristics. *Journal of Advanced Transportation* 49 (4): 535–552.
14. Yin, S. C., R. M. Xu, and Y. Zhang, et al. 2012. Signal Control Strategies for Bottleneck Area on Urban Expressway. *Journal of Transportation Systems Engineering and Information Technology* 12 (2):27–33.

A Novel Planning of Vest-Pocket Park in Historic Urban Area in Metropolis: A Case Study of Beijing

Yue Hu

Abstract It is difficult to build new parks in historic urban area in metropolis because of its complex properties: multifunction, super-high density of population and historic buildings, and conflicts between conservation and utilization of historic heritages. This study focused on exploring a novel planning method of vest-pocket parks in historic urban area in metropolis in order to solve the dilemma. The case study was conducted in the Shichahai area, Beijing, China. The information of the study area, such as existing parks, population, and historic conservative elements, was developed into GIS database which became the basis of further suitability evaluation. Each factor database layer was given a weight in the evaluation concerning their influencing degree to the final planning scheme. The research results have shown that a hierarchical vest-pocket park system with proper linkages between parks can effectively increase quantity of parks and can also function as a whole to meet the requirements of different groups of people in metropolis.

Keywords Planning · Vest-pocket park · Historic urban area · Metropolis

Introduction

Many metropolises developed from historic cities, such as Beijing, Seoul, and Rome. With the passage of time and intense development of urban in modern society, the historic urban areas play a crucial role in modern urban, and these areas have been becoming increasingly invaluable historic landscape area, traditional residential area, and flourishing commercial area in metropolises. However, the parks in these areas have some thorny problems today; it is just because historic

Y. Hu (✉)
Beijing Municipal Bureau of Landscape and Forestry, Beijing 100029, China
e-mail: yuehu1989@outlook.com

urban areas were built long time ago, especially the quantity and quality of historic urban areas are out of the line with fast-paced modern urban development.

Thanks to the complex multifunction characteristics, high density of population and buildings, and conflicts between conservation and utilization toward historic heritages, it is challenging to create new open spaces and parks. Although a large number of cities in China have increased the open space and parks in metropolises level in recent years, the increase of gross area of parks depended largely on the increase in peri-urban and suburban areas. Nevertheless, the increasing rate was extremely slow in the historic city areas. For example, in Beijing, the capital of China, from 2011 to 2015, the gross area of park increased from 19,728 to 29,502.95 ha with a growth rate of 49.55%; the park area in historic area increased only from 1022.41 to 1108.84 ha with a growth rate of 8.45%.

Vest-pocket park originated in America in 1960s. It was built in city center, where the population and buildings are in high density, in order to supply recreational places for citizens [1]. Vest-pocket park has several advantages: small land occupation, daily accessibility, and arrangement flexibility; while at the same time, it is functional limited because of its small size. There were some comprehensive studies of designing guidelines of single vest-pocket parks [2]. There were also applications of vest-pocket park in several countries and cities around the world later. In Italy, the vest-pocket parks were built along the road network which radiated from the center of the cities. In Japan, where earthquake occurs frequently, the vest-pocket parks were built for disaster prevention. In London, England, the vest-pocket parks were linked by paths to make them function as a whole systematic network. In northern Europe, there were some studies focused on the psychological restorative function of vest-pocket parks [3]. In China, the research in vest-pocket park can be dated back to 1989. There were several progresses focused on its conception, properties, case studies, and methods of single park designing. Some researchers aware of the application value of vest-pocket parks in historic urban area, but the researches focused more on the general method and less on practical planning and designing. Experience and specific method of application of vest-pocket park in historic city area are limited for people to refer.

This study was conducted in a typical historic city area, Shichahai area in Beijing, China, and focused on the planning method and process of vest-pocket parks in this area. First of all, the planning research will utilize the advantages, small land occupation and arrangement flexibility, of vest-pocket parks to increase the quality and quantity of parks. Second, it will aim at improving functional system of vest-pocket parks planned in whole study area in order to overcome the shortage of functional limitation of single vest-pocket parks. Finally, the conflicts between conservation and utilization were taken into consideration to the planning process. In our research, the planning process includes 3 main steps: (1) analyzing the site, (2) finding problems of existing parks and corresponding solutions, and (3) planning appropriate design scheme.

Study Area

The selected area was Shichahai area (Fig. 1), one of the most representative historic urban areas in Beijing, China. It is located in old city, the center of Beijing. The total area of the study area is 301.57 ha. There are 33.46 ha' water area, 144 streets, approximately 1271 valuable old trees, 40,000 families, and 10,000 inhabitants. The commercial is flourishing, and transportation is convenient. This area has already undergone historic and social changes for about 7 centuries and retained rich natural landscape resource, humanity culture, and vital historic heritages of China. Many traditional elements, such as hutong (the typical street in Beijing), siheyuan (a classical architecture style of residential housing in Beijing), and Beijing lifestyle, attract tourists from all over the world come to visit this area. There are also several important visual landscape corridors of the old city.

Data Analyses

Analyses of Existing Parks

There were 34 existing public parks (Fig. 2) available for the citizens in this area, which were mainly located near the water body and main streets. All of the 34 parks were investigated by field research and satellite-photograph observation. The existing parks were mapped by GIS technique. Then, an accessibility analysis (Fig. 3) was made by using the network analysis function of GIS. The service radius was set by 300 m according to the park's service radius standard in historic urban area published in China [4, 5]. By applying GIS techniques, the service area was defined by actual walking distance rather than the straight-line distance, which would be far more accurate than traditional circular service area.
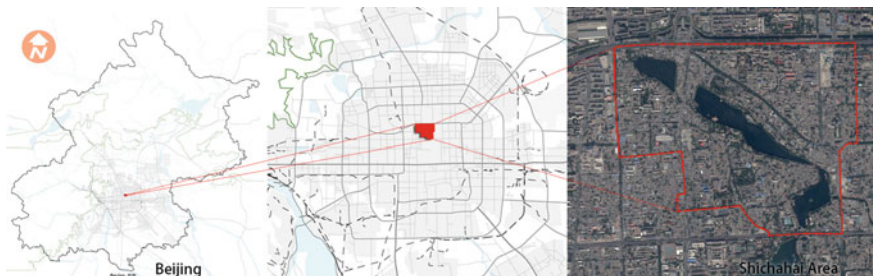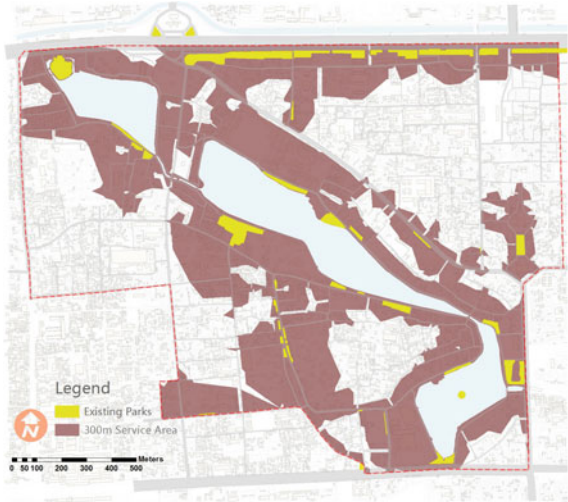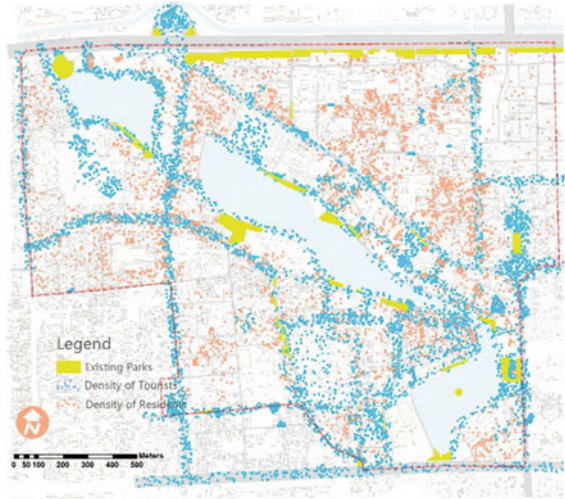

Fig. 1 Shichahai area and its location

Fig. 2 Existing parks in Shichahai area

Fig. 3 300-m service area of existing parks

Fig. 4 Population density
and distribution

The analyses showed that the existing parks were relatively poor in quantity and quality. Also, the 300 m radius' service area of all the existing parks cannot cover the whole Shichahai area.

Analyses of Population Density

Residents and tourists are two main groups consisting the population in Shichahai. The residents' density and tourists' density (Fig. 4) were mapped, respectively, according to the statistical information. Then, these two maps were overlapped to get a new map which demonstrated three different types of region regarding the group distribution: The area mainly contains residents, the area mainly contains tourists, and the area contains both residents and tourists. Each area is classified into 4 classes according to population density.

Comparing the population distribution with the 300-m service area of existing parks, it is obvious that the service area still cannot cover the regions where mainly contains residents. Besides, some streets along the popular tourism routes do not have enough parks. Moreover, although the service area covered some tourist-distributed region, the existing parks were overloaded to serve large number of visitors.

Analyses of Conservation and Utilization

There are plenty of historic and cultural heritages in Shichahai area; as a result, it is a non-negligible factor for planners to take into consideration in the planning process.

(1) Existing conservation planning

Conservation planning of the 25 historic areas in Beijing old city (conservation planning) [5] was approved and has been implemented since 2002. Shichahai is one of the 25 historic areas in the conservation planning. According to the *Conservation Planning*, Shichahai area was divided into key conservation area and surrounding regulation area (Fig. 5). Different degree of conservative management has been conducted in each area: All existing buildings in key conservation area must be protected; buildings of historic heritage and buildings of preservation must be maintained in surrounding regulation area.

(2) Historic heritage

There are a large number of historic heritages in Shichahai area, including different conservative levels. The cultural relics and historic heritage are limitation factors in planning process on one hand, and they are also opportunities to combine with planning of new parks. They can be utilized in the planning to improve landscape quality, emphasize value of heritage, revitalize historic facts, and demonstrate culture of Beijing. The historic heritage was mapped and classified into different types according to their historic value (Fig. 6).
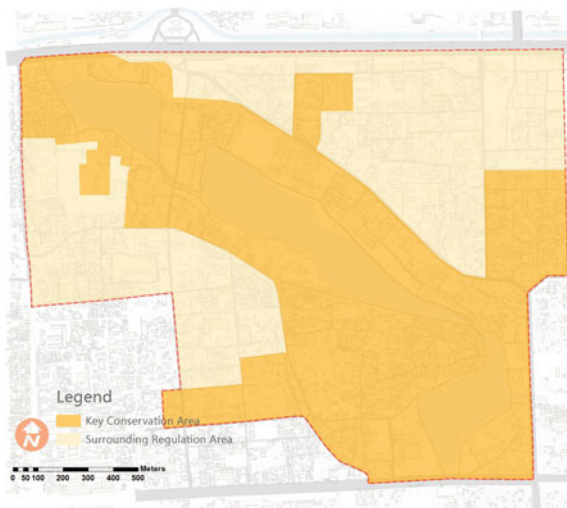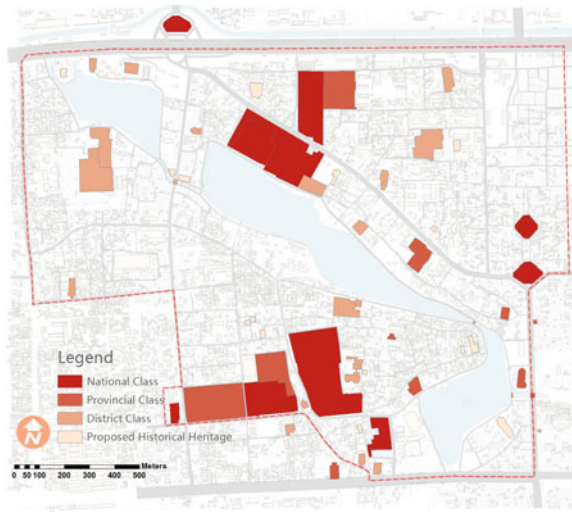
Fig. 5 Conservation areas

Fig. 6 Heritage distribution

(3) Preservation and regeneration of buildings

Buildings in Shichahai area were evaluated and classified into six types in *Conservation Planning*: buildings of historic heritage, buildings of preservation, buildings need rehabilitation, buildings can be temporarily remained, buildings need to be renovated, and buildings along main street which need to be rehabilitated in façade [5]. In the planning process, the all of the buildings in key conservation area, buildings of historic heritage, and buildings of preservation in surrounding regulation area must be carefully protected. Other buildings can be utilized to combine with newly planned vest-pocket parks. All of the buildings in Shichahai area were mapped and classified according to their own types

(4) Invaluable historic trees

Invaluable historic trees are advantageous resources distributed in Shichahai area. They can be taken into consideration in the planning process, namely, to become components of newly created vest-pocket parks. In this way, they could be protected at the same time. The historic trees were mapped, and then, the gird analyses function of GIS was applied to get the historic trees density map.

(5) Spatial structure of landscape

Shichahai area has many important landscape landmarks and corridors of old city of Beijing. In the planning process, these points and corridors can be emphasized and protected by creating more vest-pocket parks along the corridors or near the landmarks.

A 3-class hierarchical spatial landscape structure can be defined according to the scale and importance of historic heritages in Shichahai area. Each class has a

landscape center to control other landscape. The vest-pocket parks could be planned according to the spatial landscape structure in order to not only coordinate with the city's look but also function in a hierarchical way.

Problems and Solutions

Problems of the Site

(1) Limitation factors of the site

Shichahai area is an historic conservation area. The historic heritage, landscape, city texture, spatial scale, and buildings should be protected. Thus, there was very little land for building new parks.

(2) Scarcity of parks

Although there was large area of water body, the existing parks were not enough. The 300-m service area of the existing parks cannot cover the whole shichahai area. Besides, they cannot meet the requirements of high-dense population.

(3) Disorder of existing parks

The existing parks were arranged without proper order in the area and lack of hierarchy. The parks are only functioning as single small ones; they cannot work together to establish a functional system network to overcome the functional limitation shortage of small parks. Besides, the existing parks did not coordinate well with surrounding landscape structure.

Solutions and Planning Guidelines

(1) Systematic complex

Although the vest-pocket parks have limited area and scatter in the region, the planning should still make efforts to arrange them in a systematic way by establishing functional mechanism, classifying hierarchical level, and strengthening linkages between single parks.

a. Recreational system

First, vest-pocket parks should be arranged in the area which contains high-dense tourists in order to meet the requirements of the visitors. They can act as two or three of the following functions: rest area, service center, tourism-route guide, history information media, and local-lifestyle communication area.

Second, vest-pocket parks should be arranged in the area which contains high-dense residents to meet the requirements of local inhabitants. They can use them for daily exercise, communication, and having rests.

Third, tourist–resident compatible vest-pocket parks should be arranged in the area which contains both tourists and local residents in order to meet the requirements of both of the two groups of people. Perhaps, there are potential conflicts between the tourists and residents when they do different activities in this kind of parks at the same time, but this may also be an opportunity to develop this kind of parks into an interactive communication place.

b. Landscape system

Vest-pocket parks can be arranged according to spatial landscape structure of the city. The parks can be planned in the areas where landscape landmarks and the visual corridors are of significant importance to establish a harmony city look and emphasize the cityscape.

c. Hierarchical structure

A three-level vest-pocket park structure would be suitable for Shichahai area. The recreational system and landscape system should also be divided into three levels corresponding to the gross structure.

The residential recreational system can be divided into three levels: Shichahai regional level, community level, and street level. The comprehensive parks, community activity parks, and daily exercising parks can be arranged into corresponding levels, respectively.

The tourism recreational system can be also divided into three levels: sightseeing tour level, experiencing tour level, and in-depth tour level to serve tourist with different goal and purpose and guide them to the visiting place they want. If some vest-pocket parks are arranged around important attracting points, the scale and area should be expanded to serve as a tourism center toward the large number of visitors. The vest-pocket parks should be arranged along some key tourism route to guide the visitors, such as introducing the attracting points, history, and local culture.

The landscape system can be divided into three levels: city controlling landscape, regional controlling landscape, and block controlling landscape.

All the sub-system can be merged into a whole vest-pocket park system with three hierarchical levels.

d. Network linkage

The planned single linkage vest-pocket park should be properly linked by greenways along the streets to help the parks work as a whole system.

(2) Accessibility

The service radius of the vest-pocket parks should be set as 300 m according to *Evaluation Standard for Unban Landscaping and Greening* which was approved and implemented in China. In this way, citizens can get into a paring within 300 m

walking distance, which would be convenient and more equitable for the people. The service area of the planned parks should cover all or almost all of the Shichahai area. In this study, GIS was applied into the service area analyses and evaluation to get a more accurate service area.

(3) Conservation

Shichahai area is an historic urban area. The invaluable historic and cultural elements should be conserved during the planning process. Do not just consider increasing the quantity of parks but damaging the conservative elements eventually.

Methodology and Planning Scheme

Methodology

Planning According to Population Distribution Feature

Tourists and residents are two main groups of people in the area. The vest-pocket parks should be created according to the population distribution feature to meet the requirements of most of the people in the area.
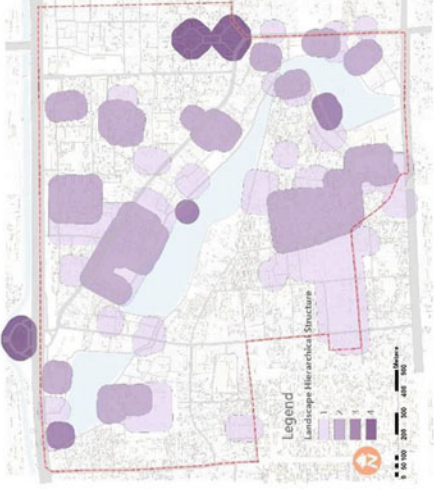
Suitability Evaluation

The information and data of Shichahai area were further processed based on GIS to get single factor evaluation map layers. These map layers were grouped into 2 kinds: summation-factor layers and subtraction-factor layers. The former group of factors are the ones that are suitable to be utilized to develop vest-pocket parks in the planning process, and the latter group of factors are the elements and regions must be protected, namely, these regions cannot be developed into vest-pocket parks. The single factor map layers were given weight values and overlapped together to generate an overall suitability evaluation map which would be the direct guide for planning.

(1) Summation-factor evaluation map layers
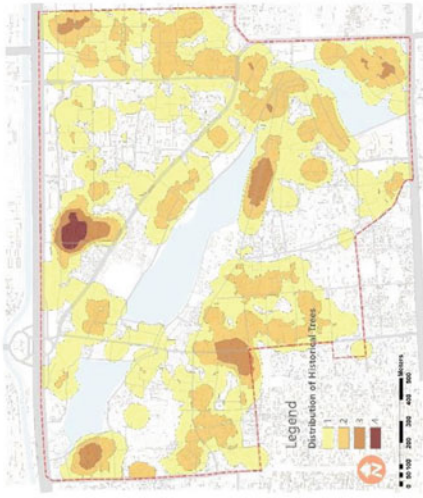
The summation-factor evaluation map layers included 6 single factor evaluation map layers: (1) spatial landscape hierarchical structure evaluation map, (2) population density evaluation map, (3) landscape corridor evaluation map, (4) tourism attractions' importance evaluation, (5) historic trees' distribution evaluation map, and (6) existing parks' service area evaluation map. Each factor was given a weight value: 0.3, 0.2, 0.2, 0.1, 0.1, and 0.1. Two of these map layers are demonstrated in the following table (Table 1) as examples:

Table 1 Examples of summation-factor evaluation map layer
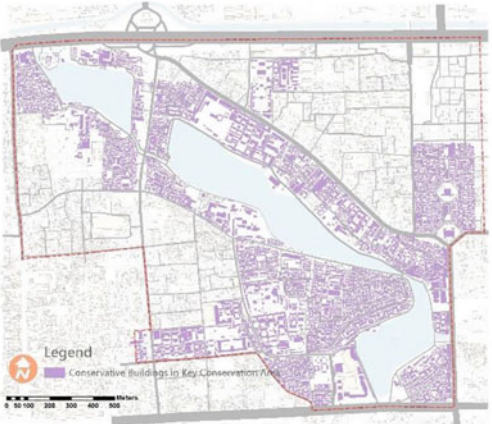
Title	Map	Explanation	Score range	Weight (%)
Spatial landscape hierarchical structure evaluation map		The 50-m-buffer of landmarks were created to show the area near these visual attracting points. Then, these areas were given scores, respectively, according to their level in the spatial landscape structure. This layer was used to control the planned parks in a systematic and hierarchical arrangement	1–4	30

(continued)

Table 1 (continued)

Title	Map	Explanation	Score range	Weight (%)
Historic trees' distribution evaluation map		This map shows the distribution of invaluable trees. This factor can be used to arrange parks around the historic trees. On one hand, these trees can improve the image of the parks. On the other hand, these trees can be protected well by setting parks around them	1–4	10

Table 2 Example of subtraction-factor evaluation map layer

Title	Map	Explanation
Map of conservative buildings in key conservation area		The map shows the must-have conservation buildings in key conservation area, which are mentioned in the *Conservation Planning*. This factor can ensure the conservation buildings avoid from being destroyed

(2) Subtraction-factor evaluation map layers

The subtraction-factor evaluation map layers included 4 single factor evaluation maps: (1) map of conservative buildings in key conservation area, (2) map of conservative buildings in surrounding regulation area, (3) map of historic heritages, and (4) map of administration and public service area. All these areas should be strictly protected during the planning process. In other words, these areas should be subtracted when overlapping all the factor map layers. One of these map layers is demonstrated in the following table (Table 2) as an example:

(3) Overlapping

All of the 10 map layers were overlapped to generate a suitability evaluation map (Fig. 7).

Planning Scheme

(1) Location layout

Suitable place for creating new vest-pocket parks was found based on the suitable evaluation (Fig. 8). Common suitable place for vest-pocket parks was discussed in people places: design guidelines for urban open spaces: street corner, crossing block, and the middle of block [6]. In this planning process, the location of vest-pocket parks was chosen based on analyses and field studies. In the real

Fig. 7 Suitability evaluation map
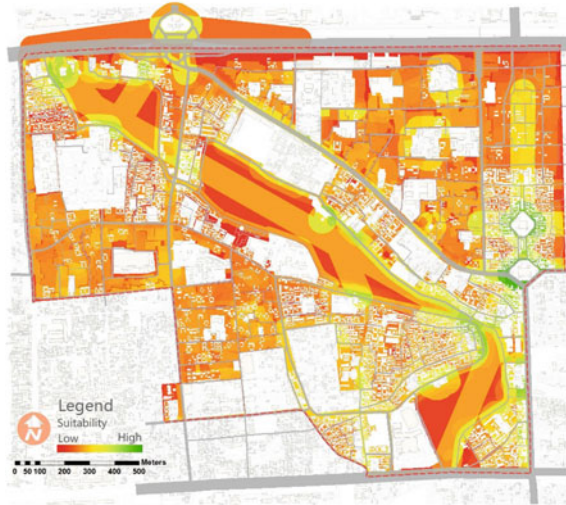
Fig. 8 Planning of location
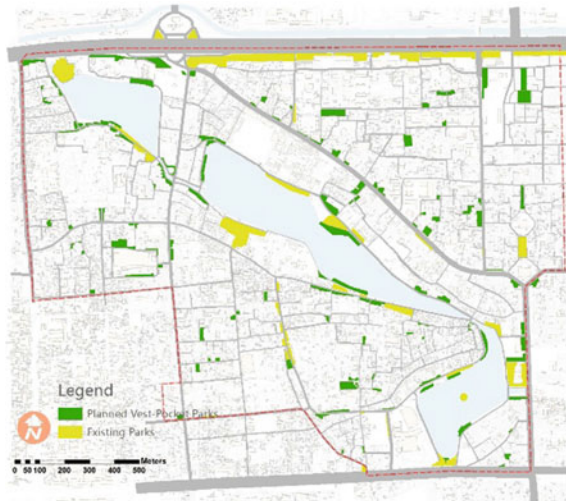
practice of planning, public participation is also recommended. Planners can select some proposed location and supply them to the residents, tourists, administrators, and planning team to discuss to perfect the planning. The planning scheme increased the quantity of parks from 34 to 165 and increased the area of parks form 83,137.76 to 164,575.47 m^2.

(2) Hierarchical classification

The vest-pocket parks were classified into 3 levels according to their corresponding suitable evaluation scores (Fig. 9).

Fig. 9 Planning of hierarchical structure
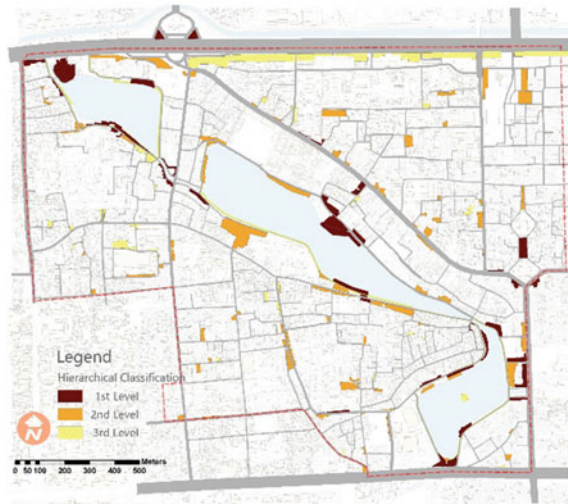
(3) Service Target

The service target of each planned vest-pocket park was decided based on the population distribution feature which was analyzed before (Fig. 10). The new parks would aim at supply different service to different group of people. For example, in the regions which contain high-dense residents, the vest-pocket parks can be designed into children's playground, fitness park, center of community activity and so on. In the regions which are popular for tourists, the parks can be designed into

Fig. 10 Planning of service target
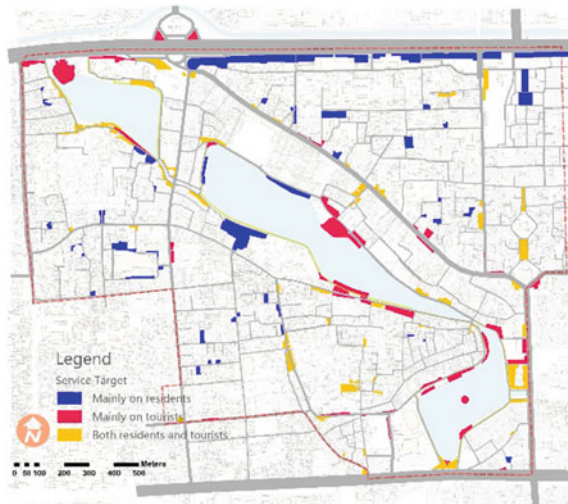
tourism service center, resting area, and historic and cultural exhibiting garden. In the regions which contain both high-dense tourists and residents, the parks can be built as folk theme park. The residents can do traditional local activities in these parks. At the same time, it is an opportunity to display the culture and history of Beijing to the visitors.

(4) Function type

Since the area of vest-pocket parks is limited, a single park can only act as one to three types of function. Thus, it is important to decide the function of each park to meet the different needs of people in different regions. Assign sightseeing and tourist center function in the higher-level parks; arrange community and daily activity function in the area which mainly contains residents; arrange religious cultural function near the temples; arrange folk-custom activity function in the area which contains both residents and tourists and so on (Fig. 11).

(5) Upgrading

Some existing parks still have some deficiencies, which need to be upgrade to become the part of vest-pocket park system (Fig. 12).

(6) Community administration

The vest-pocket parks were assigned to different communities according to location to make them more easily be administrated and maintained (Fig. 13). The community administrators can make statistical records of the way the parks are used by the citizens, which would help to improve the vest-pocket park system. Besides, the administrators can release diverse community activity information in the parks though internet and cell phones to citizens to take part in. Moreover, the administrators can organize the residents to trim the parks, which not only strengthen the

Fig. 11 Planning of function types
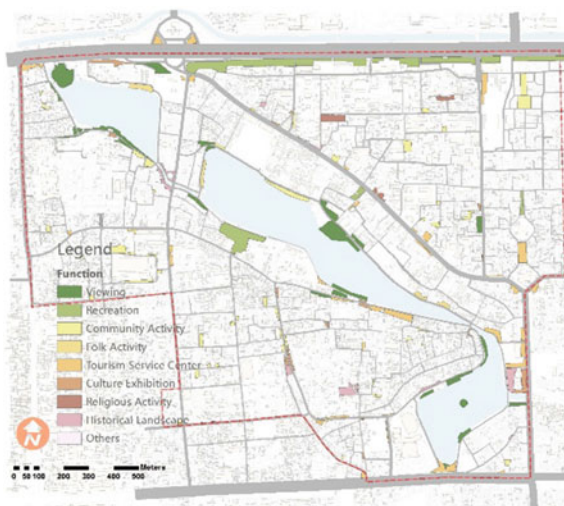
Fig. 12 Planning of upgrading existing parks
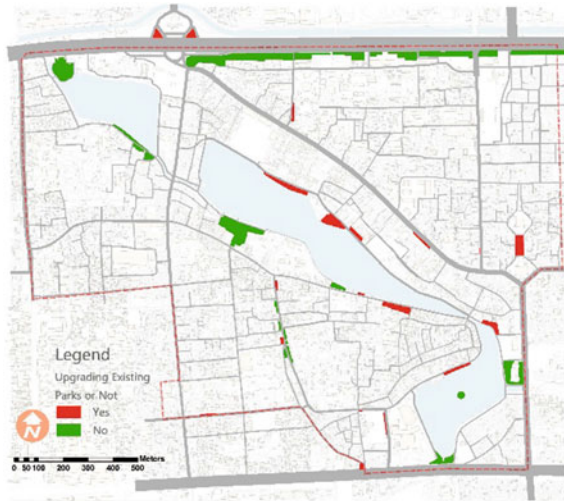
Fig. 13 Planning of community administration
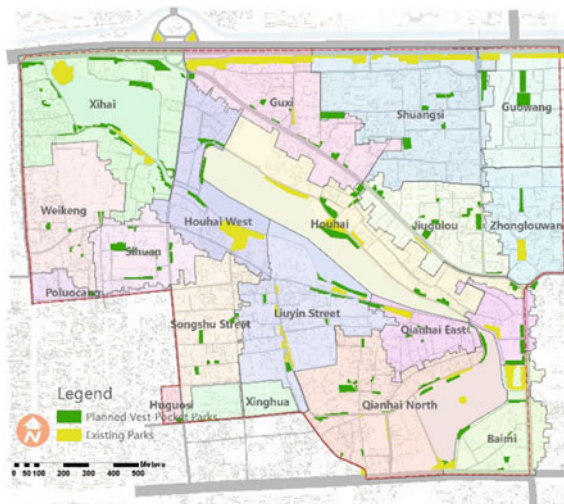
awareness of maintaining and protecting parks among residents but also lower the maintenance costs.

(7) Network linkage

The planned vest-pocket parks were linked by greenways to form a hierarchical system (Fig. 14). The level of greenways was decided by the level of linked parks. The connection of dispersed parks was strengthened by adding these linkages, which made the parks more easily accessed.

Fig. 14 Planning of network
linkage

Conclusion

In this study, the vest-pocket park system was established based on fully consideration of the problems of the site and the properties of vest-pocket parks. The planning would solve the dilemma that the parks are hard to create in historic area.

The problems of the site were defined based on comprehensive analyses of the area, including different aspect: existing condition of existing parks, population density and distribution, and conservation and utilization of historic elements. The main problems of the site are: disorder and scarcity of existing parks and the area has limitation factors hindering creation of new parks. The solution focused on establishing a hierarchical and systematic vest-pocket park network, 300-m service area coverage, and conservation of historic landscape and spatial landscape structure. The planning scheme was generated based on the suitability evaluation which was derived from overlapping influencing factor map layers. The planning scheme mainly included the location, hierarchy, function, existing-park upgrading, network linkage of the vest-pocket park system.

As to technical means, the study utilized the function of GIS, such as network analyses, overlap analyses, proximity analyses, and grid analyses, to get evaluation of the site and useful data to guide the planning.

The study shows that the deficiency and scarcity of parks in Shichahai area can be solved by adopting vest-pocket parks into this area. The application of vest-pocket parks in historic urban area in metropolis is promising. Since the area of vest-pocket parks is quite small, functional limited, the planning process should pay attention to establish a systematic vest-pocket park system which contains functional sub-system, hierarchical class, and proper network linkage.

The cutting point of this planning is the population distribution and density of two main groups of people, residents and tourists, in Shichahai area. Thus, the vest-pocket parks designed in this study focused more on recreational function which meet the main needs of people. However, the vest-pocket park system can also be planned mainly focusing on other functions, such as ecological function, disaster-prevention function, and cultural and historic education function, according to the main problem and needs of the historic urban area and the context of metropolis, which can be further studied later.

References

1. Tate, A. 2015. *Great City Parks*. Oxon: Routledge.
2. Forsyth, A., and L. Musacchio. 2005. *Designing Small Parks: A Manual for Addressing Social and Ecological Concern*. New Jersey: Wiley.
3. Nordh, H., and Østby, K. 2013. Pocket parks for people—A study of park design and use. *Urban Forestry & Urban Greening*: 12–17.
4. Ministry of Housing and Urban-Rural Development of the People's Republic of China. 2010. *Evaluation Standard for Urban Landscaping and Greening*. Beijing: China Architecture & Building Press.
5. Beijing Municipal City Planning Commission. 2002. *Conservation Planning of 25 Historic Area in Beijing Old City*. Beijing: Beijing Yanshan Press.
6. Marcus, C.C., and C. Francis. 1998. *People Places: Design Guidelines for Urban Open Space*. New Jersey: WileyInc.
7. Ministry of Housing and Urban-Rural Development of the People's Republic of China. 2005. *Code of Conservation Planning for Historic Cities*. Beijing: China Architecture & Building Press.

The Optimization of Intersection Signal in the Situation of Data Loss

Weiwei Guo, Chunling Xu, Jiyuan Tan and Leqi Li

Abstract There is a problem of intersection detector data missing in many cities because of various reasons, such as construction damage, line failure, raining, snowing or fogging, and processing errors, which will inevitably bring adverse effects on traffic flow data analysis as well as deep data mining. As a result, the signal timing plan at the intersection will be unreasonable which will exacerbate the problem of traffic jam. The intersection of Pingan Road and Xingfu Street was taken as an example, which has a critical defect problem of data missing. Experiment and analysis were made, and tireless efforts were tried to implement assumption and completion from the minimum to maximum for the southern traffic data. What's more, based on the traditional timing model calculation of Webster and HCM, the initial optimization plan can be drawn, and then, the total delay was taken as the evaluation index. Furthermore, comparison between the suggested timing plan with verification and data processing via VISSIM software and current timing plan will also be carried out. Based on this, simulation experiment in different timing plan and flow will also be carried out, and the result will be matrix-arranged. Finally, a timing plan with relatively small delay for the whole intersection was hoped to be find out no matter the southern direction traffic flow and without field investigation.

Keywords Data loss · Traffic control · Signal optimization · Delay

W. Guo · C. Xu · J. Tan (✉)
Beijing Key Lab of Urban Intelligent Traffic Control Technology,
North China University of Technology, Beijing 10014, China
e-mail: tjyphilip@163.com

L. Li
Chaowai Street Offices of Chaoyang District, Beijing 100020, China

Introduction

For a variety of reasons such as damage of construction, wiring troubles, severe weather like rain, snow or fog, and handling errors, data loss occurs frequently to the traffic data detected by detectors at urban intersections. The problem mentioned above has an impact on the perception of traffic flow and deep data analysis, which leads to the difficulty about real-time optimization of signal timing of urban intersections [1]. According to the statistics, in the traffic data of Alberta, Canada, for seven years, there were nearly half of the data facing the problem of data loss, even 90% is the ratio sometime [2]. About 40% of data had lost in the traffic data provided by the traffic administrative department of Minnesota [3]. In the process of short-term traffic flow prediction using ARIMA model and exponential smoothing model, it was found that about 20% of the testing data had lost [4]. This paper studies the methods how to optimizing timing plan of intersections under the condition of data loss, to reduce the negative effects brought by it, so as to improve the level of traffic intelligent management and transport efficiency.

At present, there is little research about the optimization of signal timing in intersections under the condition of data loss [5, 6]. The existing researches mainly focus on how to complement the missing data [7]. There are two kinds of methods to deal with the missing data: One is deterministic method and another is random method. The specific methods are as follows:

(1) Listwise deletion. It means that the records including missing value are deleted before analysis [8]. The method is relatively reasonable for the statistical analysis of data missing in a missing completely at random (MCAR) way, but deviation may exist in the processing result.
(2) Pairwise deletion. It means that make no analysis to the variable containing lost data [9].
(3) Average value. It means that replace the missing data with the average of historical data observed or others [10].
(4) Multiple linear regressions. It means that estimate the missing value using multiple linear regression algorithm [11].
(5) Expectation maximization (EM). It means that estimate the missing value based on the multiple iterations of the known data and other input variable [12]. The EM algorithm is mainly used to estimate the parameters of the probability density function of the incomplete sample [13]. The purpose of the algorithm is to maximize the expectation. The convergence rate of this algorithm is related to the loss rate of the data. Convergence rate is inversely proportional to the loss rate.
(6) Data augmentation (DA). Just like EM, the purpose of DA is looking for a maximum likelihood estimation (MLE) of the parameters, which were set according to the data set of data observed. The interpolation values generated by EM are usually used as the initial value of the missing data in the DA method [14].

However, the methods mentioned above have certain limitations to process missing data, such as the reliance on historical data and the limitation of data rate [15]. These limitations make it difficult for these methods to adapt to the variable of actual situation in the application. The purpose of this paper is to find a method to optimize the signal timing, which has universal applicability and no dependence on the historical data. Thus, it is realizable to obtain the optimal signal timing scheme on the premise of no site investigation.

Intersection Situation and Data Collection

Intersection Situation

Figure 1 is a schematic diagram that instructs the present channelization scheme of a signal intersection. Figure 2 is a schematic diagram that instructs the present signal phase of the intersection in Fig. 1. As shown in Fig. 3, there are two existing timing schemes: (1) a timing scheme for traffic flat with a cycle of 60 s and (2) another timing scheme for traffic peak with a cycle of 80 s.

Traffic Acquisition and Processing

Traffic data with no data missing were collected by detectors for three periods: morning peak, evening peak, and flat peak (Table 1).

The total traffic flow ratio should be no larger than 0.9 in order to facilitate the application of optimized timing scheme generally. According to the two kinds of classical timing methods, Webster and HCM, the total traffic flow ratio will exceed 0.9 when the traffic flow of south approach is 700 pcu/h. Therefore, assume that the

Fig. 1 Intersection channelization scheme at present
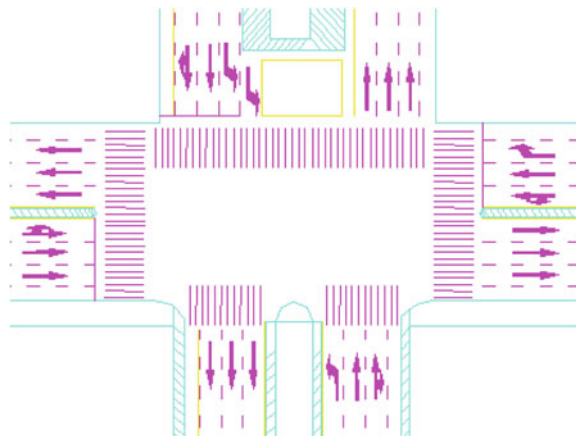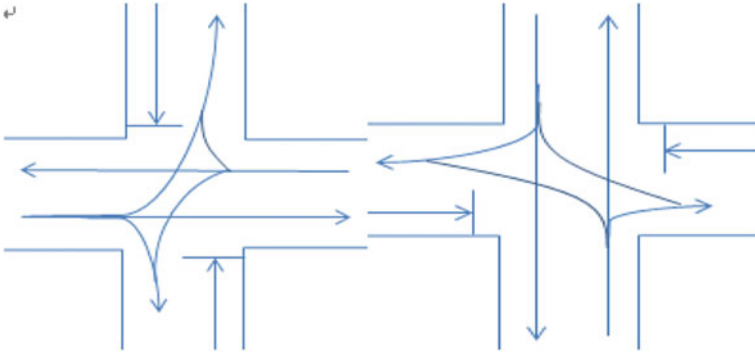
Fig. 2 Intersection signal phase at present. **a** Phase 1. **b** Phase 2

Fig. 3 Intersection signal timing diagram at present. **a** Scheme 1. **b** Scheme 2

Table 1 Traffic flow data for three periods

时段交 通量 (pcu/h)	Directions (pcu/h)											
	North			West			South			East		
	L	S	R	L	S	R	L	S	R	L	S	R
Morning peak	20	342	36	62	897	59	/	/	/	41	573	37
	398			1018			/			651		
Flat peak	22	265	23	47	760	50	/	/	/	40	676	52
	310			857			/			768		
Evening peak	26	266	25	40	869	60	/	/	/	37	670	52
	317			969			/			759		

L left; *S* straight; *R* right

traffic flow of south approach increases from 200 pcu/h to 600 pcu/h by the step 100 pcu/h. The allocation proportion of traffic flow among left, straight, and right is 1:8:1.

Optimization of Intersection Signal Timing

There are many important performance parameters for signal timing of intersections. Average vehicle delay was selected as the evaluation index of signal timing, which has been widely used at present. Next, get the signal timing scheme with

Webster and HCM. Only timing plan of morning peak was listed because of the limitation of space (Tables 2 and 3).

The Simulation and Analysis

Simulations were made using the traffic flow data of morning peak, flat peak, and evening peak calculated by two methods (Tables 4, 5 and 6).

Table 2 Timing plan of morning peak with improved Webster

Traffic (pcu/h)	Phase 1			Phase 2			Cycle (s)
	G (s)	R (s)	Y (s)	G (s)	R (s)	Y (s)	
200	17	3	3	44	3	3	67
300	17	3	3	44	3	3	67
400	17	3	3	44	3	3	67
500	56	3	3	27	3	3	89
600	85	3	3	49	3	3	140

G green; *R* red; *Y* yellow

Table 3 Timing plan of morning peak with improved HCM

Traffic (pcu/h)	Phase 1			Phase 2			Cycle (s)
	G (s)	R (s)	Y (s)	G (s)	R (s)	Y (s)	
200	6	3	3	17	3	3	29
300	6	3	3	17	3	3	29
400	6	3	3	17	3	3	29
500	21	3	3	11	3	3	38
600	34	3	3	20	3	3	60

G green; *R* red; *Y* yellow

Table 4 Comparison of average vehicle delay of morning peak

Delay	200	300	400	500	600
Now	13.1	13.2	13.5	14.6	16.3
Webster	16.7	16.9	15.9	14.5	23.1
HCM	9.6	9.7	8.9	8.1	12.9

Table 5 Comparison of average vehicle delay of flat peak

Delay	200	300	400	500	600	700
Now	10.3	10	10.1	11.7	13.4	14.3
Webster	11.7	11.2	10.9	13.3	17.2	22.3
HCM	6.9	6.4	6.7	8	9.7	12.8

It is obvious that the results of HCM were better than the others by comparing the above results. Further simulations were conducted with the data calculated by HCM next. It was uncertain that which timing scheme corresponding to 200–600 should be adopted, because there was no estimation of the lost data. For example, in the signal timing scheme, if the 'traffic' is '200,' the 'traffic' is likely to arbitrary value in actual situations. Therefore, all possible traffics between 200 and 600 were simulated to obtain the scheme with the smallest delay.

(1) Morning peak

According to the result in Table 7, the timing scheme optimized by HCM has the best effect when the traffic is 500 pcu/h. At the same time, the total intersection delay is the smallest when the traffic is 500 pcu/h. That is, no matter how much data were lost in the south approach, it is able to optimize the timing scheme with HCM, in which the traffic is 500 pcu/h (Table 8).

(2) Flat peak

According to the result in Table 9, the timing scheme optimized by HCM has the best effect when the traffic is 200 pcu/h and 300 pcu/h. At the same time, the total intersection delay is the smallest when the traffic is 200 pcu/h and 300 pcu/h. That is, no matter how much data were lost in the south approach, it is able to optimize

Table 6 Comparison of average vehicle delay of evening peak

Delay	200	300	400	500	600
Now	13.2	12.9	13.2	14.5	16.1
Webster	6.7	7	8.1	12.7	19.1
HCM	3.9	3.8	5	7.7	12.7

Table 7 Optimization of average vehicle delay of morning peak

	Timing 200	Timing 300	Timing 400	Timing 500	Timing 600
200	9.6	9.6	9.6	5.9	8.6
300	9.6	9.6	9.6	5.9	8.4
400	8.9	8.9	8.9	6.3	9
500	10.3	10.3	10.3	8.1	10.9
600	11.4	11.4	11.4	9.8	12.9
Average	9.96	9.96	9.96	7.2	9.96

Table 8 Timing scheme of morning peak with improved HCM

Phase 1			Phase 2			Cycle (s)
G (s)	R (s)	Y (s)	G (s)	R (s)	Y (s)	
1	3	3	11	3	3	38

G green; R red; Y yellow

Table 9 Optimization of average vehicle delay of flat peak

	Timing 200	Timing 300	Timing 400	Timing 500	Timing 600	Timing 700
200	6.9	6.9	7.3	7.2	7.2	8.9
300	6.4	6.4	6.8	6.7	6.9	8.6
400	6.3	6.3	6.7	6.6	6.9	8.6
500	7.6	7.6	8	8	8.3	10
600	9.1	9.1	9.4	9.4	9.7	11.6
700	9.8	9.8	10.2	10.3	10.7	12.8
Average	7.68	7.68	8.07	8.03	8.28	10.08

Table 10 Timing scheme of flat peak with improved HCM

Phase 1			Phase 2			Cycle (s)
G (s)	R (s)	Y (s)	G (s)	R (s)	Y (s)	
3	3	3	8	3	3	17

G green; *R* red; *Y* yellow

the timing scheme with HCM, in which the traffic is 200 pcu/h and 300 pcu/h (Table 10).

(3) Evening peak

According to the result in Table 9, the timing scheme optimized by HCM has the best effect when the traffic is 200 pcu/h and 300 pcu/h. At the same time, the total intersection delay is the smallest when the traffic is 200 pcu/h and 300 pcu/h. That is, no matter how much data were lost in the south approach, it is able to optimize the timing scheme with HCM, in which the traffic is 200 pcu/h and 300 pcu/h (Tables 11 and 12).

Table 11 Optimization of average vehicle delay of evening peak

	Timing 200	Timing 300	Timing 400	Timing 500	Timing 600
200	3.9	3.9	4.7	5.6	10.3
300	3.8	3.8	4.7	5.7	10.4
400	4.3	4.3	5	5.9	10
500	6.2	6.2	6.7	7.7	11.3
600	7.9	7.9	8.4	9.2	12.7
Average	5.22	5.22	5.9	6.82	10.94

Table 12 Timing scheme of evening peak with improved HCM

Phase 1			Phase 2			Cycle (s)
G (s)	R (s)	Y (s)	G (s)	R (s)	Y (s)	
11	3	3	4	3	3	21

Conclusions

The research object of this chapter is signal intersection. The detector in south approach had been damaged, and there were no data detected from it. First, find all the minimal value and the maximum value of the lost data in the south approach and complete it hypothetically. Second, missing data change from minimum to maximum, during which the preliminary scheme of intersection was calculated by Webster and HCM models. Different schemes were simulated and analyzed in the VISSIM software and were valued with the average vehicle delay index. Thus, the preliminary scheme was determined. Third, another simulation was conducted using the enumeration of traffic data, and the model has been determined. The timing plan whose delay was the smallest was the optimizing plan. The method mentioned in this chapter is suitable for the situations that are not easy to artificial investigation or that there are no historical data in an approach at all.

There are also some questions that need further research due to the limitation of time. Only one index is selected in this chapter. Indexes such as queuing length, traffic capacity, exhaust gas emission, and fuel consumption have not been taken into consideration, neither do the road condition and environment factors. Comprehensive consideration will be used for further studies. Only two traditional methods, Webster and HCM, were used to calculate the timing plan. More methods will be used for further comparison and analysis.

References

1. Chen, C., J. Kwon, and J. Rice. 2003. Detecting errors and imputing missing data for single-loop surveillance systems. Transportation Research Record, 82nd Annual Meeting of the Transportation-Research-Board, Washington, DC, 1993(1855), 160–167.
2. Zhong, M., S. Sharma, and Z.B. Liu. 2005.Assessing robustness of imputation models based on data from different jurisdictions—Examples of Alberta and Saskatchewan Canada. Transportation Research Record, 84th Annual Meeting of the Transportation-Research-Board, Washington, D.C. No. 1917, 16–126.
3. Yuwei, B., W. Xu, and Q.Z. Tony. 2016. Online method to impute missing loop detector data for urban freeway traffic control. *Transportation Research Record* 2593: 37–46.
4. Haizhong, W., L. Lu, and D. Shangjia. 2016. A novel work zone short-term vehicle-type specific traffic speed prediction model through the hybrid EMD-ARIMA framework. *Transportmetrica B-Transport Dynamics* 4 (3): 159–186.
5. Jinjun, T., Z. Guohui, and W. Yinhai. 2015. A hybrid approach to integrate fuzzy C-means based imputation method with genetic algorithm for missing traffic volume data estimation. *Transportation Research Part C-Emerging Technologies* 51: 29–40.
6. Kristian, H., Z. Yajie, and W. Yinhai. 2015. Flexible and robust method for missing loop detector data imputation. *Transportation Research Record* 2527: 29–36.

7. Jie, L., H.J. van Zuylen, and Guorong Wei. 2014. Diagnosing and interpolating loop detector data errors with probe vehicle data. *Transportation Research Record* 2423: 61–67.
8. Krishnamoorthy, K. 2013. Comparison of confidence intervals for correlation coefficients based on incomplete monotone samples and those based on listwise deletion. *Journal of Multivariate Analysis* 114: 378–388.
9. Al-Sakib Khan, P., M. Muhammad Mostafa, and J. Jinfang. 2011. An efficient approach of secure group association management in densely deployed heterogeneous distributed sensor network. *Security and Communication Networks* 4(9): 1013–1026.
10. Dongfang, M., F. Fengjie, and J. Sheng. 2016. Recognition of critical links for gating using queue detector data. *Ksce Journal of Civil Engineering* 20 (7): 2955–2964.
11. Jinjun, T., Z. Yajie, and A. John. 2016. Travel time estimation using freeway point detector data based on evolving fuzzy neural inference system. *Plos One* 11(2).
12. Gurcan, C., B. Anton, and C. Mecit. 2016. Adaptive traffic parameter prediction: Effect of number of states and transferability of models. *Transportation Research Part C-Emerging Technologies* 72: 202–224.
13. Yingjie, X., S. Xingmin, and S. Guanghua. 2016. Towards improving quality of video-based vehicle counting method for traffic flow estimation. *Signal Processing* 120: 672–681.
14. Nianfeng, W., V. Ardalan, and L. Andre. 2016. Reconstructing maximum likelihood trajectory of probe vehicles between sparse updates. *Transportation Research Part C-Emerging Technologies* 65: 16–30.
15. Huachun, T., W. Yuankai, and C. Bin. 2014. Robust missing traffic flow imputation considering nonnegativity and road capacity. *Mathematical Problems in Engineering*. doi: 10.1155/2014/763469