Advances in Intelligent Systems and Computing 538

Masato Akagi
Thanh-Thuy Nguyen
Duc-Thai Vu
Trung-Nghia Phung
Van-Nam Huynh *Editors*

Advances in Information and Communication Technology

Proceedings of the International Conference, ICTA 2016

Springer

Advances in Intelligent Systems and Computing

Volume 538

Series editor

Janusz Kacprzyk, Polish Academy of Sciences, Warsaw, Poland
e-mail: kacprzyk@ibspan.waw.pl

About this Series

The series "Advances in Intelligent Systems and Computing" contains publications on theory, applications, and design methods of Intelligent Systems and Intelligent Computing. Virtually all disciplines such as engineering, natural sciences, computer and information science, ICT, economics, business, e-commerce, environment, healthcare, life science are covered. The list of topics spans all the areas of modern intelligent systems and computing.

The publications within "Advances in Intelligent Systems and Computing" are primarily textbooks and proceedings of important conferences, symposia and congresses. They cover significant recent developments in the field, both of a foundational and applicable character. An important characteristic feature of the series is the short publication time and world-wide distribution. This permits a rapid and broad dissemination of research results.

Advisory Board

Chairman

Nikhil R. Pal, Indian Statistical Institute, Kolkata, India
e-mail: nikhil@isical.ac.in

Members

Rafael Bello, Universidad Central "Marta Abreu" de Las Villas, Santa Clara, Cuba
e-mail: rbellop@uclv.edu.cu

Emilio S. Corchado, University of Salamanca, Salamanca, Spain
e-mail: escorchado@usal.es

Hani Hagras, University of Essex, Colchester, UK
e-mail: hani@essex.ac.uk

László T. Kóczy, Széchenyi István University, Győr, Hungary
e-mail: koczy@sze.hu

Vladik Kreinovich, University of Texas at El Paso, El Paso, USA
e-mail: vladik@utep.edu

Chin-Teng Lin, National Chiao Tung University, Hsinchu, Taiwan
e-mail: ctlin@mail.nctu.edu.tw

Jie Lu, University of Technology, Sydney, Australia
e-mail: Jie.Lu@uts.edu.au

Patricia Melin, Tijuana Institute of Technology, Tijuana, Mexico
e-mail: epmelin@hafsamx.org

Nadia Nedjah, State University of Rio de Janeiro, Rio de Janeiro, Brazil
e-mail: nadia@eng.uerj.br

Ngoc Thanh Nguyen, Wroclaw University of Technology, Wroclaw, Poland
e-mail: Ngoc-Thanh.Nguyen@pwr.edu.pl

Jun Wang, The Chinese University of Hong Kong, Shatin, Hong Kong
e-mail: jwang@mae.cuhk.edu.hk

More information about this series at http://www.springer.com/series/11156

Masato Akagi · Thanh-Thuy Nguyen
Duc-Thai Vu · Trung-Nghia Phung
Van-Nam Huynh
Editors

Advances in Information and Communication Technology

Proceedings of the International Conference, ICTA 2016

Springer

Editors
Masato Akagi
School of Information Science,
Area of Human Life Design
Japan Advanced Institute of Science
and Technology
Nomi-shi, Ishikawa
Japan

Thanh-Thuy Nguyen
Department of Computer Science
VNU University of Engineering
and Technology
Hanoi
Vietnam

Duc-Thai Vu
Faculty of Information Technology
Thai Nguyen University of Information
and Communication Technology
Thai Nguyen
Vietnam

Trung-Nghia Phung
Thai Nguyen University of Information
and Communication Technology
Thai Nguyen
Vietnam

Van-Nam Huynh
School of Knowledge Science,
Area of Knowledge Management
Japan Advanced Institute of Science
and Technology
Nomi-shi, Ishikawa
Japan

ISSN 2194-5357 ISSN 2194-5365 (electronic)
Advances in Intelligent Systems and Computing
ISBN 978-3-319-49072-4 ISBN 978-3-319-49073-1 (eBook)
DOI 10.1007/978-3-319-49073-1

Library of Congress Control Number: 2016955568

© Springer International Publishing AG 2017
This work is subject to copyright. All rights are reserved by the Publisher, whether the whole or part of the material is concerned, specifically the rights of translation, reprinting, reuse of illustrations, recitation, broadcasting, reproduction on microfilms or in any other physical way, and transmission or information storage and retrieval, electronic adaptation, computer software, or by similar or dissimilar methodology now known or hereafter developed.
The use of general descriptive names, registered names, trademarks, service marks, etc. in this publication does not imply, even in the absence of a specific statement, that such names are exempt from the relevant protective laws and regulations and therefore free for general use.
The publisher, the authors and the editors are safe to assume that the advice and information in this book are believed to be true and accurate at the date of publication. Neither the publisher nor the authors or the editors give a warranty, express or implied, with respect to the material contained herein or for any errors or omissions that may have been made.

Printed on acid-free paper

This Springer imprint is published by Springer Nature
The registered company is Springer International Publishing AG
The registered company address is: Gewerbestrasse 11, 6330 Cham, Switzerland

Preface

This volume contains the papers presented at the International Conference on Advances in Information and Communication Technology (ICTA2016), which was held in Thai Nguyen city, Vietnam, during 12–13 December 2016. The conference was jointly organized by Thai Nguyen University of Information and Communication Technology (ICTU), Institute of Information Technology—Vietnam Academy of Science (IoIT), Fengchia University, Taiwan (FCU), Japan Advanced Institute of Science and Technology (JAIST) and National Chung Cheng University, Taiwan (CCU). The principal aim of ICTA2016 Conference is to bring together researchers, academics, practitioners and students in order to not only share research results and practical applications but also to foster collaboration in research and education in information and communication technology. The ICTA2016's Program Committee received a total of 150 submissions. Each submission was peer reviewed by at least two members of the Program Committee. Finally, 66 papers were chosen for presentation at Conference Sections and publication in the proceedings. Besides the main track, the conference has three invited speeches. We would like to express our appreciation to all the members of the Program Committee for their support and cooperation in this publication. We would like to thank Prof. Janusz Kacprzyk (Series Editor) and Dr. Thomas Ditzinger (Executive Editor, Interdisciplinary and Applied Science, Engineering, Springer) for their support and cooperation in this publication. We are also thankful to Anand Chozhan and his colleagues at Springer for providing a meticulous service for the timely production of this volume. Last but not the least, we wish to thank all the authors and participants for their contributions and fruitful discussions that made this conference a success.

December 2016 ICTA 2016

Masato Akagi
Thanh-Thuy Nguyen
Duc-Thai Vu
Program Chairs

Organization

Honorary Chairs

Bach Hung Khang	IoIT, Vietnam
Pham Viet Binh	ICTU, Vietnam

General Chairs

Nguyen Van Tao	ICTU, Vietnam
Nguyen Huu Cong	TNU, Vietnam

Program Chairs

Masato Akagi	JAIST, Japan
Nguyen Thanh-Thuy	VNU-UET, Vietnam
Vu Duc-Thai	LICTU, Vietnam

Publication Chairs

Huynh Van-Nam	JAIST, Japan
Phung Trung-Nghia	ICTU, Vietnam
Tran Dang Hung	Hanoi National University of Education, Vietnam
Nguyen Van Toi	ICTU, Vietnam

Organizing Chairs

Dang Thi Oanh	ICTU, Vietnam
Dam Thanh Phuong	ICTU, Vietnam
Le Hung Linh	ICTU, Vietnam
Ngo Manh Tuong	ICTU, Vietnam

Nguyen Hai Minh	ICTU, Vietnam
Nguyen Thanh Ha	TNU, Vietnam
Nguyen Van Nui	TNU, Vietnam
Nguyen Truong Thang	IoIT, Vietnam

Program Committee

Bui Thu Lam, Vietnam
Bui Ngoc Thang, Vietnam
Bui Quang Anh, Vietnam
Ching-Hwa Cheng, Taiwan
Chow Yen Desmond Sim, Taiwan
Chih-Hung Chou, Taiwan
Chyi-Ren Dow, Taiwan
Dang Quang A, Vietnam
Dang The Ngoc, Vietnam
Dang Thi Oanh, Vietnam
Dang Thanh Phu, Vietnam
Dang Van Duc, Vietnam
Dao Huy Du, Vietnam
Dao Phuong Nam, Vietnam
Dao Trung Kien, Vietnam
Do Nang Toan, Vietnam
Do Thanh Nghi, Vietnam
Do Thi Bac, Vietnam
Do Thi Ngoc Diep, Vietnam
Do Thi Tu Anh, Vietnam
Doan Nhat Quang, Vietnam
Doan Van Ban, Vietnam
Don-Gey Liu, Taiwan
Duong Chinh Cuong, Vietnam
Fang-Rong Hsu, Taiwan
Julia Tzu-Ya Weng, Taiwan
Gwo-Jiun Horng, Taiwan
Ha Duyen Trung, Vietnam
Ha Tien Ngoan, Vietnam
Hamido Fujita, Japan
Heng-Yi Su, Taiwan
Ho Dang Phuc, Vietnam
Ho Van Huong, Vietnam
Hoang Do Thanh Tung, Vietnam
Hoang Manh Thang, Vietnam
Huynh Van-Nam, Japan
Huynh Xuan Hiep, Vietnam
Kun-Chih Chen, Taiwan
Lai Hien-Phuong, Vietnam
Lai Khac Lai, Vietnam
Le Ba Dung, Vietnam
Le Duc Hau, Vietnam
Le Dung Muu, Vietnam
Le Hai Chau, Vietnam
Le Hoai Bac, Vietnam
Le Hoang Son, Vietnam
Le Hung Linh, Vietnam
Le Huy Tung, Vietnam
Le Quang Minh, Vietnam
Le Ngoc Tu, Japan
Le Nhat Thang, Vietnam
Le Thanh Ha, Vietnam
Le Thi Lan, Vietnam
Le Thi My Hanh, Vietnam
Le Trong Vinh, Vietnam
Le Trung Thanh, Vietnam
Le Viet Phuong, France
Luong Chi Mai, Vietnam
Luong The Dung, Vietnam
Masato Akagi, Japan
Nghiem Thi Phuong, Vietnam
Ngo Duc Thien, Vietnam
Ngo Hong Son, Vietnam
Ngo Quoc Tao, Vietnam
Ngo Thi Duyen, Vietnam
Ngo Xuan Bach, Vietnam
Nguyen Ai Viet, Vietnam
Nguyen Cat Ho, Vietnam
Nguyen Chien Trinh, Vietnam
Nguyen Cong Dieu, Vietnam
Nguyen Cuong, Vietnam
Nguyen Dai Hai, Vietnam

Nguyen Dinh Thuan, Vietnam
Nguyen Duc Dung, Vietnam
Nguyen Duc Nhan, Vietnam
Nguyen Duc Thuan, Vietnam
Nguyen Duy Minh, Vietnam
Nguyen Gia Nhu, Vietnam
Nguyen Hai Duong, Vietnam
Nguyen Hai Minh, Vietnam
Nguyen Hoai Nam, Vietnam
Nguyen Hoang Anh, Vietnam
Nguyen Hoang Nam, Vietnam
Nguyen Hoang Phuong, Vietnam
Nguyen Hong Quang, Vietnam
Nguyen Huu Cong, Vietnam
Nguyen Huu Quynh, Vietnam
Nguyen Kim Anh, Vietnam
Nguyen Linh Trung, Vietnam
Nguyen Long Giang, Vietnam
Nguyen Minh Tuan, Vietnam
Nguyen Nhi Gia Vinh, Vietnam
Nguyen Ngoc Linh, Vietnam
Nguyen Phuong Huy, Vietnam
Nguyen Quoc Cuong, Vietnam
Nguyen Quoc Khuong, Vietnam
Nguyen Quy Hy, Vietnam
Nguyen Son, Vietnam
Nguyen Tai Hung, Vietnam
Nguyen Thai Ha, Vietnam
Nguyen Thanh Hai, Vietnam
Nguyen Thanh-Thuy, Vietnam
Nguyen The Quang, Vietnam
Nguyen Thi Oanh, Vietnam
Nguyen Thi Phuong Giang, Vietnam
Nguyen Thi Thu Ha, Vietnam
Nguyen Thi Thuy, Vietnam
Nguyen Toan Thang, Vietnam
Nguyen Tran Hiep, Vietnam
Nguyen Truong Thang, Vietnam
Nguyen Van Chi, Vietnam
Nguyen Van Huan, Vietnam
Nguyen Van Huy, Vietnam
Nguyen Van Nui, Vietnam
Nguyen Van Tam, Vietnam
Nguyen Van Tao, Vietnam
Nguyen Van Thuy, Vietnam
Nguyen Van Toi, Vietnam
Nguyen Van Vy, Vietnam
Nguyen Viet Anh, Vietnam
Nguyen Viet Dung, Vietnam
Nguyen Viet Hung, Vietnam
Nguyen Xuan Huy, Vietnam
Nong Thi Hoa, Vietnam
Phan Thuong Cang, Vietnam
Pham Duc Long, Vietnam
Pham Minh Trien, Vietnam
Pham Ngoc Hung, Vietnam
Pham Quang Nhat Minh, Japan
Pham Thanh Giang, Vietnam
Pham Thanh Nam, Taiwan
Pham Viet Binh, Vietnam
Pham Xuan Nghia, Vietnam
Phan Xuan Hieu, Vietnam
Phung Trung-Nghia, Vietnam
Rakshtith Mysore Rajashekar, UK
Ren-Hung Hwang, Taiwan
Ta Tuyen, Vietnam
Tay-Jyi Lin, Taiwan
Teh-Lu Liao, Taiwan
Thai Quang Vinh, Vietnam
Tien-Yin Jimmy Chou, Taiwan
To Van Khanh, Vietnam
Tran Dang Hung, Vietnam
Tran Dinh Que, Vietnam
Tran Duc Tan, Vietnam
Tran Hoai Linh, Vietnam
Tran Hoang Tung, Vietnam
Tran Quy Nam, Vietnam
Tran Thanh Hai, Vietnam
Tran Thi Ngan, Vietnam
Tran Thu Ha, Vietnam
Tran Trung Duy, Vietnam
Tran Van Lang, Vietnam
Tran Xuan Nam, Vietnam
Truong Ha Hai, Vietnam
Truong Kien, USA
Truong Quoc Dinh, Vietnam
Vu Chan Hung, Vietnam
Vu Chien Thang, Vietnam
Vu Duc-Thai, Vietnam
Vu Duc Thi, Vietnam

Vu Duy Hai, Vietnam
Vu Duy Linh, Vietnam
Vu Duy Loi, Vietnam
Vu Hoai Linh, Vietnam
Vu Hoang Giang, Vietnam
Vu Manh Xuan, Vietnam
Vu Nhu Lan, Vietnam
Vu Tat Thang, Vietnam
Vu Tuyet Trinh, Vietnam
Vu Viet Vu, Vietnam
Vu Vinh Quang, Vietnam
Yao-Ting Huang, Taiwan
Yao-Tung Tsou, Taiwan
Yi-Chung Chen, Taiwan

Contents

Keynote Addresses

Multimodal Based Clouds Computing Systems for Healthcare and Risk Forecasting Based on Subjective Analysis 3
Hamido Fujita

Telematics and Advanced Transportation Services 5
Chyi-Ren Dow

Toward Affective Speech-to-Speech Translation 6
Masato Akagi

ICTA 2016 Main Track

A Computer Vision Based Machine for Walnuts Sorting Using Robot Operating System 9
Truong Tran, Tu Nguyen, Mai Nguyen and Tuan Pham

A FPGA Based Two Level Optimized Local Filter Design for High Speed Image Processing Applications 19
Majida Kazmi, Arshad Aziz, Pervez Akhtar and Nassar Ikram

A Frequency Dependent Investigation of Complex Shear Modulus Estimation 31
Quang Hai Luong, Manh Cuong Nguyen and Tran Duc Tan

A Method to Enhance the Remote Sensing Images Based on the Local Approach Using KMeans Algorithm 41
Trung Nguyen Tu, Duc Dang Van, Huy Ngo Hoang and Thoa Vu Van

A Method for Clustering and Identifying HTTP Automated Software Communication 53
Manh Cong Tran, Hai Nam Nguyen, Minh Hieu Nguyen and Yasuhiro Nakamura

A New Neuro-Fuzzy Inference System for Insurance Forecasting 63
Le Hoang Son, Mai Ngoc Khuong and Tran Manh Tuan

A New Schema to Identify S-farnesyl Cysteine Prenylation Sites with Substrate Motifs 73
Van-Nui Nguyen, Thi-Xuan Tran, Hai-Minh Nguyen, Hong-Tan Nguyen and Tzong-Yi Lee

A Novel Framework Based on Deep Learning and Unmanned Aerial Vehicles to Assess the Quality of Rice Fields 84
Nguyen Cao Tri, Tran Van Hoai, Hieu N. Duong, Nguyen Thanh Trong, Vo Van Vinh and Vaclav Snasel

A Semi-supervised Learning Method for Hybrid Filtering 94
Thi Lien Do and Duy Phuong Nguyen

A Study on Fitness Representation in Genetic Programming 104
Thuong Pham Thi, Xuan Hoai Nguyen and Tri Thanh Nguyen

Adaptive Robust Ability of High Order Sliding Mode Control for a 3-D Overhead Crane System 113
Dao Phuong Nam, Nguyen Doan Phuoc and Nguyen Thi Viet Huong

An Evaluation of Hand Pyramid Structure for Hand Representation Based on Kernels 123
Van-Toi Nguyen, Thi-Lan Le and Thanh-Hai Tran

An Evolutionary-Based Term Reduction Approach to Bilingual Clustering of Malay-English Corpora 132
Rayner Alfred, Leow Ching Leong and Joe Henry Obit

An Exploratory Study on Students' Performance Classification Using Hybrid of Decision Tree and Naïve Bayes Approaches 142
Yoong Yen Chuan, Wahidah Husain and Amirah Mohamed Shahiri

An Improved Method for Stock Market Forecasting Combining High-Order Time-Variant Fuzzy Logical Relationship Groups and Particle Swam Optimization 153
Nghiem Van Tinh and Nguyen Cong Dieu

An Iterative Method to Solve Boundary Value Problems with Irregular Boundary Conditions 167
Vu Vinh Quang and Truong Ha Hai

BKCA, an E-Consultancy System for Studying 178
Thi Thu Trang Nguyen, Viet Thang Dinh and Tuan Dung Cao

Classifying Human Body Postures by a Support Vector Machine with Two Simple Features 188
Nguyen Van Tao, Nong Thi Hoa and Quach Xuan Truong

Cluster Analysis, Classification and Forecasting Tool on DS30 for Better Investment Decision 197
Arshiful Islam Shadman, Sheikh Shadab Towqir, M. Ahnaf Akif, Mehrab Imtiaz and Rashedur M. Rahman

Comparing Modified PSO Algorithms for MRS in Unknown Environment Exploration 207
Anh-Quy Hoang and Minh-Trien Pham

Design Adaptive-CTC Controller for Tracking Target Used Mobile Robot-Pan Tilt-Stereo Camera System 217
Le Van Chung and Duong Chinh Cuong

Estimation Localization in Wireless Sensor Network Based on Multi-objective Grey Wolf Optimizer 228
Trong-The Nguyen, Ho Thi Huong Thom and Thi-Kien Dao

Evaluation of Mobile Phone Wireless Charging System Using Solar and Inductive Coupling 238
Norizam Sulaiman, Mohd Shawal Jadin, Amran Abdul Hadi, Muhamad Sharfi Najib, Mohd Ashraf Ahmad and Nurul Shazrien Amsyha Yusuf

Hardware Implementation of MFCC Feature Extraction for Speech Recognition on FPGA 248
Van-Lan Dao, Van-Danh Nguyen, Hai-Duong Nguyen and Van-Phuc Hoang

Improved Adaptive Fuzzy Sliding Mode Control for Second Order Nonlinear System 255
Pham Van Thiem, Lai Khac Lai and Nguyen Thi Thanh Quynh

Improved LDPC Iterative Decoding Algorithm Based on the Reliable Extrinsic Information and Its Histogram 265
Nguyen Anh Tuan and Pham Xuan Nghia

Improving Control Quality for Two-Wheeled Mobile Robot 275
Gia Thi Dinh, Nguyen Duy Cuong and Nguyen Dang Hao

Incorporation of Experience and Reference-Based Topic Trust with Interests in Social Network 286
Phuong Thanh Pham, Manh Hung Nguyen and Dinh Que Tran

Integrating Multisignature Scheme into the Group Signature Protocol 294
Hung Dao Tuan, Hieu Minh Nguyen, Cong Manh Tran, Hai Nam Nguyen and Moldovyan Nikolay Adreevich

Interestingnesslab: A Framework for Developing and Using Objective Interestingness Measures 302
Lan Phuong Phan, Nghia Quoc Phan, Ky Minh Nguyen, Hung Huu Huynh, Hiep Xuan Huynh and Fabrice Guillet

Inverted Pendulum Control Using Fuzzy Reasoning Method Based on Hedge Algebras by Approach to Semantic Quantifying Adjustment of Linguistic Value 312
Nguyen Duy Minh, Do Thi Mai and Nguyen Thi Thu Hien

***k*-Nearest Neighbour Using Ensemble Clustering Based on Feature Selection Approach to Learning Relational Data** 322
Rayner Alfred, Kung Ke Shin, Mohd Shamrie Sainin, Chin Kim On, Paulraj Murugesa Pandiyan and Ag Asri Ag Ibrahim

Low Power ECC Implementation on ASIC 332
Van-Lan Dao, Van-Tinh Nguyen and Van-Phuc Hoang

Malwares Classification Using Quantum Neural Network 340
Tu Tran Anh and The Dung Luong

Managing Secure Personal Mobile Health Information 347
Chan Wai Chen, Mohd Azam Osman, Zarul Fitri Zaaba and Abdullah Zawawi Talib

Mobile Online Activity Recognition System Based on Smartphone Sensors 357
Dang-Nhac Lu, Thu-Trang Nguyen, Thi-Thu-Trang Ngo, Thi-Hau Nguyen and Ha-Nam Nguyen

Multi-criteria Path Planning for Disaster Relief: An Example Using the Flood Risk Map of Shalu District, Taiwan 367
Sheng-Wei Qiu, Yi-Chung Chen, Tsu-Chiang Lei, Hsin-Ping Wang and Hsi-Min Chen

Multi-feature Based Similarity Among Entries on Media Portals 373
Thi Hoi Nguyen, Dinh Que Tran, Gia Manh Dam and Manh Hung Nguyen

MyEpiPal: Mobile Application for Managing, Monitoring and Predicting Epilepsy Patient 383
Nur Ayuni Marzuki, Wahidah Husain and Amirah Mohamed Shahiri

New Block Ciphers for Wireless Moblile Netwoks 393
Pham Manh Tuan, Bac Do Thi, Minh Nguyen Hieu and Nam Do Thanh

Numerical Method for Solving a Strongly Mixed Boundary Value Problem in an Unbounded Domain 403
Dang Quang A and Dinh Hung Tran

Numerical Solution of a Fully Fourth Order Nonlinear Problem 413
Dang Quang A and Ngo Thi Kim Quy

On the Performance of Decode-and-Forward Half-Duplex Relaying with Time Switching Based Energy Harvesting in the Condition of Hardware Impairment 421
Tan N. Nguyen, Phuong T. Tran, Huong-Giang Hoang, Hoang-Sy Nguyen and Miroslav Voznak

Opportunistic Multiple Relay Selection Schemes in both Full-Duplex and Half-Duplex Operation for Decode-and-Forward Cooperative Networks 431
Hoang-Sy Nguyen, Anh-Hoa Thi Bui, Nhat-Tan Nguyen and Miroslav Voznak

Pattern Discovery in the Financial Time Series Based on Local Trend 442
Mai Van Hoan, Dao The Huy and Luong Chi Mai

Performance Evaluation of SAR ADC with Organic Semiconductor . . . 452
Huyen Thanh Pham, Toan Thanh Dao and Thang Vu Nguyen

Performance Evaluation of Wireless Networks Based on Testbed. 460
Ngo Hai Anh, Takumi Tamura and Pham Thanh Giang

Predicting Early Crop Production by Analysing Prior Environment Factors 470
Tousif Osman, Shahreen Shahjahan Psyche, MD Rafik Kamal, Fouzia Tamanna, Farzana Haque and Rashedur M. Rahman

Prediction of Generalized Anxiety Disorder Using Particle Swarm Optimization 480
Wahidah Husain, Saw Hui Yng, Nur'Aini Abdul Rashid and Neesha Jothi

Quality Improvement of Vietnamese HMM-Based Speech Synthesis System Based on Decomposition of Naturalness and Intelligibility Using Non-negative Matrix Factorization 490
Anh-Tuan Dinh, Thanh-Son Phan and Masato Akagi

Reducing Middle Nodes Mapping Algorithm for Energy Efficiency in Network Virtualization. 500
Tran Manh Nam, Nguyen Van Huynh and Nguyen Huu Thanh

Research on Enhancing Security in Cloud Data Storage 510
Minh Le Quang, Phan Huy Anh, Nguyen Anh Chuyen and Le Khanh Duong

Resource-Aware Scheduling in Heterogeneous, Multi-core Clusters for Energy Efficiency 520
Xuan T. Tran

Secured-OFS: A Novel OpenFlow Switch Architecture with Integrated Security Functions 530
Bao Ho, Quoc Nguyen, Cuong Pham-Quoc and Tran Ngoc Thinh

Semi-supervised Clustering in Fuzzy Min-Max Neural Network 541
Dinh Minh Vu, Viet Hai Nguyen and Ba Dung Le

Smart Lecture Room for Smart Campus Building Automation System 551
Kien Tran Pham Thai

Solving Navier-Stokes Equation Using FPGA Cellular Neural Network Chip 562
Duc-Thai Vu, Le Hung Linh and Nguyen Mai Linh

Special Characters of Vietnamese Sign Language Recognition System Based on Virtual Reality Glove 572
Diep Nguyen Thi Bich, Trung-Nghia Phung, Thang Vu Tat and Lam Phi Tung

Stochastic Bounds for Inference in Topic Models 582
Xuan Bui, Tu Vu and Khoat Than

Swarm Intelligence-Based Approach for Macroscopic Scale Odor Source Localization Using Multi-robot System 593
Anh-Quy Hoang and Minh-Trien Pham

The Analyzes of Network-on-Chip Architectures Based on NOXIM Simulator 603
Van-Nam Dinh, Mau-Viet Ho, Van-Cuong Nguyen, Tung-Son Ngo and Effiong Charles

The Asynchronous Cooperative Amplify-and-Forward Relay Network with Partial Feedback to Improve the System Performance 612
The-Nghiep Tran, Van-Bien Pham and Huu-Minh Nguyen

The Optimization Model for Calculating Distribution System Planning Integrated Photovoltaic 622
V.V. Thang

The Prediction of Succinylation Site in Protein by Analyzing Amino Acid Composition 633
Van-Minh Bui and Van-Nui Nguyen

The Signal Control for the Traffic Network via Image Analysis 643
Du Dao Huy and Mui Nguyen Duc

Toward Cyber-Security Architecture Framework for Developping Countries: An Assessment Model 652
Nguyen Ai Viet, Le Quang Minh, Doan Huu Hau, Nguyen Ngoc Tuan, Nguyen Nhat Quang, Nguyen Dinh Chinh, Nguyen Van Luc and Pham Thanh Dat

Author Index 659

Keynote Addresses

Multimodal Based Clouds Computing Systems for Healthcare and Risk Forecasting Based on Subjective Analysis

Hamido Fujita(✉)

Intelligent Software Systems, Iwate Prefectural University, Takizawa, Japan
HFujita-799@acm.org

Abstract. In decision making most approaches are taking into account objective criteria, however the subjective correlation among decision makers provided as preference utility is necessary to be presented to provide confidence preference additive among decision makers reducing ambiguity and produce better utility preferences measurement for subjective criteria among decision makers. Most models in Decision support systems are assuming criteria as independent. Therefore, these models are ranking alternatives based on objective data analysis. Also, different type of data (time series, linguistic values, interval data, etc.) imposes some difficulties to do decision making using classical multi criteria decision making models.

Sophisticated machine learning methods to estimate or extract emotions from the content created by users has been developed including support vector machines, Bayesian networks, maximum entropy approaches and concept level analysis of natural language text, supported by combinations of common-sense reasoning. These approaches are mainly based on language text processing with sufficient documents, which is usually inlarge is not available. We think Subjectiveness is related to the contextual form of criteria. Uncertainty of some criteria in decision making is also considered as other important aspect These draw backs in decision making are major research challenges that are attracting wide attention, like on big data analysis for risk prediction, medical diagnosis and other applications that are in practice more subjective to user situation and its knowledge related context. Subjectivity would be examined based on correlations between different contextual structures that is reflecting the framework of personal context, for example in nearest neighbor based correlation analysis fashion. Some of the attributes incompleteness also may lead to affect the approximation accuracy. Attributes with preference-ordered domain relations properties become one aspect in ordering properties in rough approximations.

The Virtual Doctor System (VDS) developed by my group is a system assisting human doctor who is practicing medical diagnosis in real situation and environment. The interoperability is represented by utilizing the medical diagnosis cases of medical doctor, represented in machine executable fashion based on human patient interaction with virtual avatar resembling a real doctor. VDS is practiced as a virtual avatar interacting with the human patient based on physical views and mental view

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_1

analysis. In this talk I outline our VDS system and then discuss related issues in subjective decision making in medical domain. Using fuzzy reasoning techniques in VDS, it has been shown that it is possible to provide better precision in circumstances that is related to partial known data and uncertainty on the acquisition of medical symptoms.

Telematics and Advanced Transportation Services

Chyi-Ren Dow(✉)

Department of Information Engineering and Computer Science,
Feng Chia University, Taichung, Taiwan
crdow@fcu.edu.tw

Abstract. Our Telematics and Advanced Transportation Alliance, a minor alliance between academia and industry was founded by Ministry of Science and Technology, Taiwan in 2012. In order to enhance the technical ability of telematics and provide advanced and high-quality transportation services, our alliance consists of scholars and experts from the areas of telematics and advanced traffic management. The core technologies of the alliance include Controller Area Network (CAN) bus, CANopen communication technology for Industry 4.0, WAVE DSRC networks, APP hardware/software integration, big data analysis for driving safety, and fleet management services. Our alliance has more than twenty industry members so far and has held more than 80 technical seminars, promotional activities, education training programs, and factory visits. In addition to providing personnel trainings, technical services and industry guidance for our members, innovation cooperation services are highlighted to strengthen industry academic cooperation. To integrate resource and promote services, two new core technologies, CANopen communication technology for Industry 4.0 and analysis service of intersections and reasons prone to accidents, are provided to our members and we also promote the open Internet of Vehicle (IoV) platform developed by our alliance. To strengthen the industry upgrade and guidance, seven core technologies and IoV platform are used to integrate our members' products and assist them to get resources from Taiwan government. For autonomous operation and sustainable development, competitive advantages of our core technologies are used to develop new products and new technologies with our members to achieve the sustainability of our alliance.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_2

Toward Affective Speech-to-Speech Translation

Masato Akagi(✉)

Japan Advanced Institute of Science and Technology, 1-1 Asahidai, Nomi, Ishikawa 923-1292, Japan
akagi@jaist.ac.jp

Abstract. Speech-to-speech translation (S2ST) is the process by which a spoken utterance in one language is used to produce a spoken output in another language. The conventional approach to S2ST has focused on processing linguistic information only by directly translating the spoken utterance from the source language to the target language without taking into account para-linguistic and non-linguistic information such as the emotional states at play in the source language. This paper introduces activities of JAIST Acoustic Information Science Laboratory, Human Life Design Area, Japan Advanced Institute of Science and Technology that explore how to deal with para- and non-linguistic information among multiple languages, with a particular focus on speakers' emotional states, in S2ST applications called "affective S2ST." In our efforts to construct an effective system, we discuss (1) how to describe emotions in speech and how to model the perception/production of emotions and (2) the commonality and differences among multiple languages in the proposed model. We then use these discussions as context for (3) an examination of our "affective S2ST" system in operation.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_3

ICTA 2016 Main Track

A Computer Vision Based Machine for Walnuts Sorting Using Robot Operating System

Truong Tran[1(✉)], Tu Nguyen[2], Mai Nguyen[3], and Tuan Pham[2]

[1] Center for Training of Excellent Students,
Hanoi University of Science and Technology, Hanoi, Vietnam
20110889@student.hut.edu.vn
[2] Center for Optoelectronics,
National Center for Technological Progress, Hanoi, Vietnam
{ngoctu,phamhongtuan}@cfoc.vn
[3] Department of Precision Mechanical and Optical Engineering,
Hanoi University of Science and Technology, Hanoi, Vietnam
mai.nguyenthiphuong@hust.edu.vn

Abstract. Industrial machines are generally expensive to implement due to their requirement of being *fast and robust*. In this paper, we proposed a new approach to the problem, particularly, a walnut sorting machine that was built cheaply using open solution. On hardware side, we used *readily available* electronics and on the software side, *Robot Operating System* was adopted to handle low level hardware abstraction. *Machine Learning*, *Computer Vision* and data processing techniques were implemented on top of them using high level programming language. This resulted in a highly functional, easy to maintain yet inexpensive walnuts sorting machine - which was confirmed by our Austrian partner's testing. This success, hopefully, will pave the way for more projects using similar approaches.

Keywords: Robot operating system · Machine learning · Computer vision · Walnuts sorting machine

1 Introduction

Walnut contains high level of monounsaturated fatty acids, 72 % of which is polyunsaturated fatty acids. The substance is known for lowering various cardiometabolic risk factors and walnut consumption is showed to correlate with the decrease of Type 2 Diabetes in Women [1]. According to Food and Agriculture Organization of the United Nations, worldwide production of walnut has been rising rapidly in recent years. In 2012, China, the biggest producer in the world, produced 1.7 billion tons of walnut. At the same time, The United States produced merely more than 400 thousand tons of walnut, however, they were the largest exporter. Challenges in storage and transport hindered developing countries from exporting more, such as: Non-uniformity size and weight causes

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_4

problems for offering the product and food processing. Fungal mold infested walnut spread to others, spoils the batch and creates aflatoxin - a potent carcinogen, and so on. Those challenges raised a need for walnut quality checking.

In the past few years, many non-destructive inspection models have been proposed for walnut [2,3] and other fruits [4]. Using different approaches to analyze data, those studies all rely on acoustic characteristic to infer the internal quality. This is a necessity since visual cue can only give out external quality. This direction, although being novelty and innovative, is impractical in current factory scenarios. Acoustic noise is very difficult to get rid of and complicated techniques require substantial more computational power that only large-scale businesses can afford.

This project's main goal is to develop a machine with feasible and non-destructive methods to sort out walnuts. Recalled that internal quality is difficult to measure directly; we combined 3 characteristic of Visual appearance, Mass and Specific Mass to give out a highly accurate prediction for it. Since imaging sensors and loadcells are much easier to shell from background noise, this machine is more suitable for factories environment. In our design, all characteristics of walnuts are measured in the same line. They are then physically sorted into 5 categories: A, B, C, D and F (unqualified).

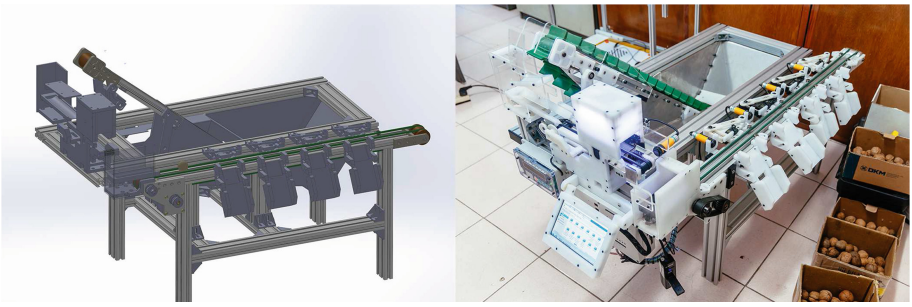


Fig. 1. CAD model and real image of the walnut sorting machine

By using open-source software and open-source friendly electronics, our design is inexpensive and very modular. Most of machinery parts are simulated in Computer Aided Design before production. This resulted in a machine that is cost effective, highly configurable and easy to make. Figure 1 shows the CAD model and our real system.

2 Materials and Methods

2.1 Materials

For the experimental works, total 100 unsorted walnuts were collected from three gardens, located in different places in Austria. Collection of walnuts were

performed in four batches, 25 nuts each time. Steps were taken to ensure the randomness of the collection process in each batch. After the collection, each nut was tagged with a unique serial number, generated on the basis of batch number and origin gardens.

Three independent human experts work in the relevant fields were selected for manual setting the *Ground Truth* of walnuts quality. Each expert was given caliper and magnifying glass for non-destructive measurement on the outside of the nut. Later, destructive tests were carried out to determine nutmeat quality. This complicated process ensured the *Ground Truth* were free (or close to free) from error.

2.2 Methods

A schematic diagram of our machine is showed in Fig. 2. Walnuts were carried by *Conveyor 1* through the line to *Conveyor 2*. Each of them was weighted by *LoadCell* and taken picture by *Camera*. Going through *Conveyor 3*, the walnut is pushed to its designated category by 4 mechanical selectors and 1 discard conveyor.

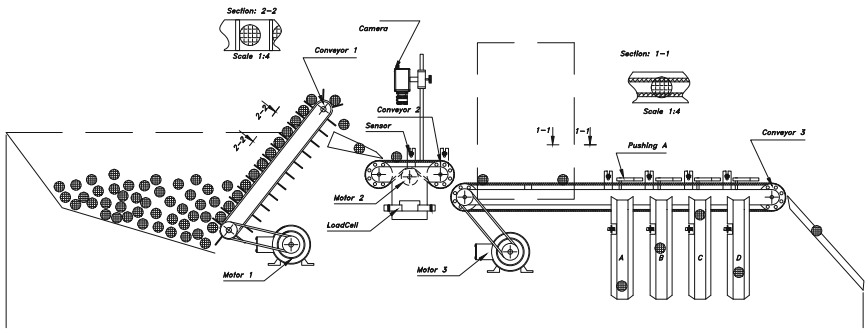


Fig. 2. Machine schematic diagram

Computer Vision. The camera had FullHD resolution and was interfaced with the mainboard through an USB 2.0 port. Using V4L2 API, we can turn on/off, change ISO, take picture from it and transmit data back and forth with the CPU. Assume a walnut has Ellipse shape, its edge co-ordinations should satisfy Eq. 1

$$ax^2 + 2bxy + cy^2 + dx + ey = 1. \tag{1}$$

From captured images, ellipse shape of a walnut is measured using Algorithm 1. First, a picture of background environment is taken. Subtract it from pictures with walnuts going through we have the walnuts' shapes. A shape itself is a two-dimensional array, we ran it through *Canny filter* [5] to compute edge co-ordinations. The resultant *Matrix(x,y)* is filtered through *findContours* function to find a boundary with ellipse shape *Matrix01(x,y)*. Using a modified

RANSAC algorithm [6], Eq. 1 can be solved by model fitting all edge points from *Matrix01(x,y)* into it, giving a, b, c, d, e.

Estimate of position angle (pp) and position angle of an undetermined axis (pa) are calculated as in Eq. 2:

$$pp = \arctan(\frac{2b}{a-c}) \tag{2a}$$

$$pa = \frac{1}{2pp} \tag{2b}$$

Axial ratio (r) is calculated as in Eq. 3:

$$r = \sqrt{\frac{1 + \frac{2b}{\sin pp(a+c)}}{1 - \frac{2b}{\sin pp(a+c)}}} \tag{3}$$

X-centre of ellipse (x_0), Y-centre of ellipse (y_0), the length of the undetermined axis (ab) and position angle in degrees (pa_d) are calculated as in Eq. 4:

$$x_0 = \frac{b \times e - c \times d}{a \times c - b^2} \tag{4a}$$

$$y_0 = \frac{b \times d - a \times e}{a \times c - b^2} \tag{4b}$$

$$ab = \sqrt{y_1(x_0^2 + \frac{1}{a}) + y_2 + y_3} \qquad \text{with} \begin{cases} y_1 &= \sin pa^2 + (r \cos pa)^2 \\ y_2 &= x_0 y_0 (r^2 - 1) \sin pp \\ y_3 &= y_0^2 (\cos pa^2 + (r \sin pa)^2) \end{cases} \tag{4c}$$

If $r < 1$, ($majorAxis$), ($minorAxis$) and position angle in degrees (pa_d) are calculated in Eq. 5:

$$majorAxis = \frac{ab}{r} \tag{5a}$$

$$minorAxis = ab \tag{5b}$$

$$pa_d = \frac{pa \times 45}{\arctan 1} - 90 \tag{5c}$$

If $r \geq 1$, ($majorAxis$), ($minorAxis$) and position angle in degrees (pa_d) are calculated in Eq. 6:

$$majorAxis = ab \tag{6a}$$

$$minorAxis = \frac{ab}{r} \tag{6b}$$

$$pa_d = \frac{pa \times 45}{\arctan 1} \tag{6c}$$

To reduce the complexity of our algorithm, we used $a = \frac{majorAxis + minorAxis}{2}$ as the dimension parameter for walnuts evaluation process.

Algorithm 1. Ellipse Parameter Extraction

```
procedure EllipseExtract
    image00 ← take picture of background
    image01 ← take picture of a walnut
    image02 ← image01 − image00
loop:
    if (x,y) ≤ size of image02 then
        Matrix(x, y) ← CannyFilter image02.
        (x,y)++.
        goto loop.
        close;
    Matrix01(x,y) ← findContours Matrix(x,y).
    a,b,c,d,e ← MODfitEllipseRANSAC Matrix01(x,y).
    Return a,b,c,d,e.
```

Fungal mold is a common disease which spread from infection. Since the nutmeat is protected by a harden shell, the shell would be infected first and turned dark. Depend on how severe the infection is; such walnuts were to be lower ranked or discarded entirely. Since walnuts' shapes and colors vary greatly among themselves, a decision was made to use Machine Learning for visual appearance classification.

Working directly with images of walnuts would require huge computational effort, which is not desirable for our project. Although the walnut can be separated efficiently from background - as mentioned in previous part - the process of detect the walnut using *sliding windows* still takes a lot of time. Recall that we only consider the overall shade of the walnut, a study of Histogram showing *brightness vs probability* such as in Fig. 3 could be sufficient. The strategy is described below:

1. Setup training data of 300 labeled walnuts (Healthy/Moderate/Bad)
2. Construct and normalized a Histogram for each of the walnuts
3. Setup parameters for SVM, number of iterations, tolerance error...etc.
4. Train the SVM with OpenCV CvSVM::train library
5. Classification with OpenCV CvSVM::predict library

SVM - *Support Vector Machines*, first proposed by Cortes & Vapnik [7] is a *Supervised Learning* technique. It means if sufficient labeled data is provided, such as in this case of 3 type shell shades, the algorithm will output an optimal hyperplane which clearly separated one from the 2 others. By compiling the histogram, we avoid the sliding windows problem (since data is normalized) and make sure the input is linearly separable (i.e one hyperplane can clearly separate A from not A groups). All those steps can be done using OpenCV library [8]. The resulting file is then used in the machine for classification of walnuts. Since the two processes are asymmetric, with the classification much less computational intense, we can implement it on a Quad core ARM Cortex-A9. Walnuts going through the sorting machine will be categorized as: Healthy (Fig. 3.a),

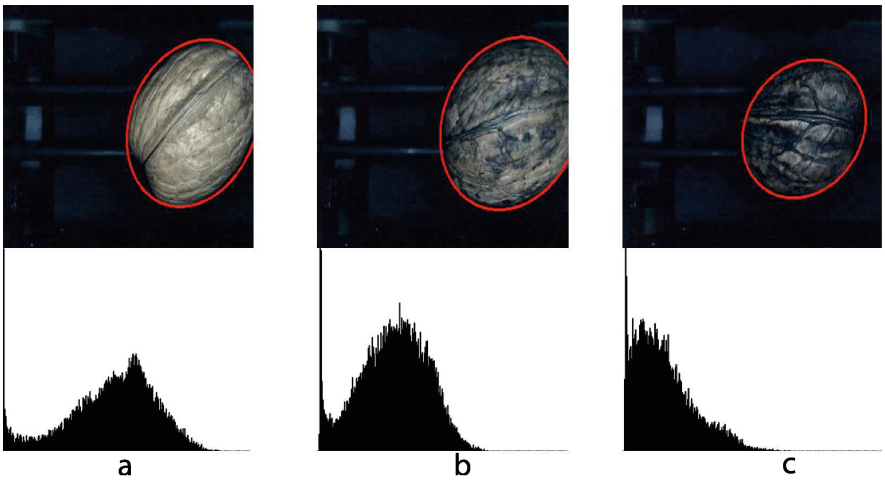


Fig. 3. Three level of shell shades and their respective histograms

Moderate (Fig. 3.b) and Bad (Fig. 3.c).This information would help sorting walnuts into five groups, which will be discussed later in Sect. 2.3, Sorting algorithm.

Loadcell Signal Processing. The *Loadcell* was placed as showed in Fig. 2, continuously measured the walnut weight while it going through *Conveyor2*. Due to the vibration from *Motor1, Motor2, Motor3* and various mechanical parts in real system in Fig. 1 the measured values fluctuated. It was challenging to denoise those values since the noise came from multi sources and was not in a repeated pattern. However, according to the *Central Limit Theorem*: if the noise was small and from multi source, it could be considered random and if sufficiently large number of trials were taken, the measured values will be approximately *normally distributed*.

Metaphorical speaking, if you hold a ball above a line and let it free fall, it may not land exactly at the line; but, if you trial enough times, the probability that it landed close to the line is greater than the probability that it landed faraway. In fact, the *Probability Distribution* peaked at the *median*, that is most of the time it landed at the line. Using this reasoning, we proposed "a hack" to measure the weight more precisely. When a walnut go though *weightSensor1*, *weightValue* is measured in succession multi times every *0.1 second* until it reached *weightSensor2*. The measured values were normalized and the median was chosen as the weight value.

Robot Operating System. (ROS)[1] - as its name implies, is an open-source Operating System specifically designed for Robotic. It provides structured communications layer above the host operating systems of a heterogeneous compute

[1] http://www.ros.org.

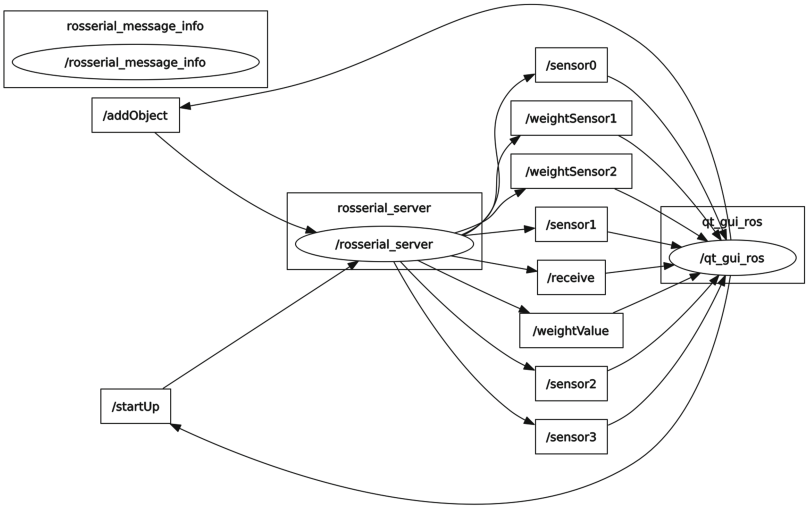


Fig. 4. Robot Operating System (ROS) - Backbone of the sorting machine

cluster. Figure 4 shows ROS as the backbone of our machine. *qt_gui_ros, rosserial_server* and other processes are connected at runtime in a peer-to-peer topology. This allowed us to design the machine as many connected modules. ROS uses microkernel design and pushes all complexity in libraries, making debug easier since components can be debugged individually.

User interacts with the walnuts sorting machine through user interface *qt_gui_ros*. When *startUp* is called, *rosserial_server* send *rosserial_message_info* to *sensor* and *weightSensor* to check their statuses. When every devices are ready, their statues are displayed in *qt_gui_ros* and *addObject* is called. It turns on *conveyor 1*, *conveyor 2*, *conveyor 3* and motors respectively. The startup process is now complete, advanced functions such as *Computer Vision* and *Load-cell signal processing* will be executed. The shutdown process is similar to start up, but in reverse. ROS has a very good track record and still grows exponentially [9]. By utilizing ROS, our walnuts sorting machine benefits from its state of the art architecture and well documented functionality.

2.3 Sorting Algorithm

The sorting algorithm is showed in Algorithm 2; with a_1 to a_4 and b_1 to b_4 are user input-able values; a is the dimension parameter (mm) while b represents weight (g). Depends on the combination of two values, T is assigned from 4 to 0, which means quality from A to F. Then, if m is greater than 115 % the max weight of next quality level, the quality is increased 1 level. Similarly, if visual appearance c = 3, T will stay the same; if c = 2, T will downgrade 1 level and if c = 1, T will downgrade to F quality regardless of its previous value.

Algorithm 2. Sorting Algorithm

procedure GROUPASSIGN
- **if** r $\leq a_1$ and m $\geq b_1$ **then** T = 1 ▷ T=1 means D quality
- **if** r $\leq a_2$ and m $\geq b_2$ **then** T = 2 ▷ T=2 means C quality
- **if** r $\leq a_3$ and m $\geq b_3$ **then** T = 3 ▷ T=3 means B quality
- **if** r $\leq a_4$ and m $\geq b_4$ **then** T = 4 ▷ T=4 means A quality
- **else**
  - T = 0 ▷ T=0 means F quality

Upgrade:
- **if** m $\geq 115\,\%\ b_{t+1}$ **then** T = T+1

Downgrade:
- **if** c = 3 **then** T = T ▷ c=3 means Healthy color
- **if** c = 2 **then** T = T-1 ▷ c=2 means Moderate color
- **if** c = 1 **then** T = 0 ▷ c=1 means Bad color

Walnuts size and shape varies greatly among different types and locations; gardens owners also want control over quality standards for different markets. Because of that, a and b are user defined and so does the relationship of c and T. In our particular case of Austria walnuts; they were set for maximize profit as: $a_1 = 34$, $a_2 = 38$, $a_3 = 41$, $a_4 = 50$, $b_1 = 6$, $b_2 = 9.6$, $b_3 = 12.6$, $b_4 = 14.6$.

3 Result and Discussion

3.1 Sensors Calibration

Using V4L2 API, camera's parameters are set as: Resolution: 640×480, Contrast: 75, Auto gain: Disabled, Gain: +195, Brightness: +155, Sharpness: +90, Auto focus: Disabled, Focus Absolute: +199, White Balance: Enabled. Loadcell's parameters are set as: Resolution: 0.01 g, Max Weight: 600 g, Min Weight: 0.2 g.

3.2 Result

Tables 1 and 2 respectively showed the performance of *Machine* and *Worker* vs the *Ground Truth*. Both of them have done a very good job of sorting walnuts into 5 categories, which indicated in the less than 2 % *coefficient of variation* (C.V). In general, the Human Worker had lower C.V than the Machine. However, at this level of accuracy both results are well acceptable. The *Average Time* to complete the task, unfortunately, is much longer for the Worker compared to the Machine (nearly 4 times), due to the Machine's inherently more efficient measuring line.

Table 1. The machine's 5 times trial. Abbreviation: C.T − Categories, G.T − Ground Truth, S.D − Standard Deviation, R.S.D − Relative Standard Deviation, C.V − Coefficient of Variation, T − Time, A.T − Average Time

C.T	G.T	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	S.D	R.S.D	C.V(%)
A	10	10	10	10	10	9	0.18	0.018	1.79
B	33	34	33	32	33	33	0.28	0.009	0.86
C	40	39	40	41	41	41	0.36	0.009	0.89
D	14	14	14	14	13	14	0.18	0.013	1.28
F	3	3	3	3	3	3	0.00	0.000	0.00
T(s)		213	197	188	220	235			
A.T(s)				211					

Table 2. Worker's 5 times trial. Abbreviation: C.T − Categories, G.T − Ground Truth, S.D − Standard Deviation, R.S.D − Relative Standard Deviation, C.V − Coefficient of Variation, T − Time, A.T − Average Time

C.T	G.T	1^{st}	2^{nd}	3^{rd}	4^{th}	5^{th}	S.D	R.S.D	C.V(%)
A	10	10	10	10	10	10	0.00	0.000	0.00
B	33	33	34	33	33	33	0.18	0.005	0.54
C	40	40	40	40	41	40	0.18	0.004	0.45
D	14	14	13	14	13	14	0.22	0.016	1.56
F	3	3	3	3	3	3	0.00	0.000	0.00
T(s)		767	813	822	779	807			
A.T(s)				798					

3.3 Discussion

At present state, we have constructed a fully functioning walnuts sorting machine based on walnuts' sizes, weights and healthy levels. *Machine Learning* and *Computer Vision* techniques are used to extract walnuts' sizes and health while algorithm is implemented to get more precise weight from loadcell sensor. The proposed system also aims to be easy to use, modular and highly accurate. Its speed is limited by conveyors and the mechanical selectors rather than the embed system, which works on the order of microsecond.

Test has been conducted only for the Austrian walnut variety but the sorting machine can be used for others with reasonably change in shape and color without modification. With a small modification, it can work with other nuts and firm skin fruits. Our machine is well shielded in critical area that its performance is not affected by variation of external factors like changing in ambient light, wind direction...etc. Our study shows that, the Machine Learning/Computer Vision based system's accuracy is close to manual worker while being much faster. Future work is to implement Computer Vision on X-ray scanning or Machine Learning on vibration characteristic of walnuts to better understand their internal quality.

Acknowledgments. We would like to thank Center for Optoelectronics, National Center for Technological Progress, Vietnam and its Austrian partner for financing the project. We also want to express our gratitude to Department of Precision Mechanical and Optical Engineering, Hanoi University of Science and Technology, Vietnam for supporting the project with technical insight.

References

1. Pan, A., Sun, Q., Manson, J.E., Willett, W.C., Hu, F.B.: Walnut consumption is associated with lower risk of type 2 diabetes in women. J. Nutr. **143**(4), 512–518 (2013)
2. Khalesi, S., Mahmoudi, A., Hosainpour, A., Alipour, A.: Detection of walnut varieties using impact acoustics and artificial neural networks (ANNs). Modern Appl. Sci. **6**(1), 43 (2011)
3. Khalifa, S., Komarizadeh, M.H., et al.: An intelligent approach based on adaptive neuro-fuzzy inference systems (ANFIS) for walnut sorting (2012)
4. Diezma-Iglesias, B., Ruiz-Altisent, M., Barreiro, P.: Detection of internal quality in seedless watermelon by acoustic impulse response. Biosyst. Eng. **88**(2), 221–230 (2004)
5. Canny, J.: A computational approach to edge detection. IEEE Trans. Pattern Anal. Mach. Intell. **8**(6), 679–698 (1986)
6. Fischler, M.A., Bolles, R.C.: Random sample consensus: a paradigm for model fitting with applications to image analysis and automated cartography. Commun. ACM **24**(6), 381–395 (1981)
7. Cortes, C., Vapnik, V.: Support-vector networks. Mach. Learn. **20**(3), 273–297 (1995)
8. Bradski, G., et al.: The opencv library. Doctor Dobbs J. **25**(11), 120–126 (2000)
9. Cousins, S.: Exponential growth of ROS [ROS topics]. IEEE Rob. Autom. Mag. **18**(1), 19–20 (2011)

A FPGA Based Two Level Optimized Local Filter Design for High Speed Image Processing Applications

Majida Kazmi(✉), Arshad Aziz, Pervez Akhtar, and Nassar Ikram

Department of Electrical Engineering (PNEC), National University of Sciences and Technology (NUST), H-12, Islamabad, Pakistan
majida.kazmi@pnec.nust.edu.pk

Abstract. This work presents an efficient Feild Programmable Gate Array (FPGA) based local filter design for portable and high speed image processing applications. It is highly optimized by using two level optimization. The first level optimization at design-level exploits tempo-spatial parallelism of filters by developing parallel/pipelined architecture. For exploiting spatial-parallelism, design computes partial results of multiple MACs in parallel and accumulates them via adder-tree for final result. Though it bears good performance aptitude but adder-tree incurs long critical path (4.713 ns) thus limits design performance. The critical path was reduced to 2.489 ns with temporal parallelism by pipelining the adder-tree. Design performance is further enhanced by deploying the second level optimization at post-implementation level where device aware floor-planning fine tunes the design. It aligns all utilized embedded resources of design on Xilinx Virtex-5 device and confines slice based logic across them. It results in packing the design within small area with reduced slice count and critical path (2.32 ns). After applying two levels of optimization, the design occupies 89 Slices, 3 DSP-Slices, 2 BRAM18 and achieves high frequency of 431.03 MHz.

Keywords: Local filters · Real time image processing · FPGA · XSG tool · Floor planning

1 Introduction

Digital image processing becomes a subject of common interest in emerging areas such as computer vision, medical, surveillance and industrial to name a few. In all of these applications, image processing operations help to improve the quality of images for its correct interpretation and analysis without human interference [1]. Most of these image processing operations are local by nature, in which calculation of an output pixel depends on a local neighborhood of input pixels. These operations are computationally expensive and time-consuming. Software based sequential architectures for such operations are relatively slow, thus not able to meet time constraints associated with real time systems [17].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_5

Therefore to accelerate computation of these image processing operations, hardware such as Field Programmable Gate Array (FPGA) based parallel architectures have been explored [1,6,17].

Among different local image processing operations, one of the most frequently used operation is two dimensional convolution in local filters [4]. In the convolution operation, output pixel P_{out} is obtained by convolving a $N \times N$ window of input pixels with the $N \times N$ Filter Kernel which generally requires N^2 multiplications and N^2-1 additions. The quadratic growth of N factor implies that it is a computationally expensive and time consuming operation [6,13,17,20]. The FPGA based parallel architectures exploit tempo-spatial parallelism of local filtering operation to accelerate its computation for real time applications [1,6,13,16,17,20]. At the same time, due to its inherent re-configurability, it is also possible to modify the filter design at any stage to incorporate requirements of any specific application. For these reasons, the FPGA platform is viable for modeling the filter designs from its development phase till its final design.

Xilinx System Generator (XSG) is a tool for modeling and rapid prototyping of the image processing systems onto FPGA. It represents image processing system on Simulink/MATLAB platform and uses High Level Descriptions (HLD) to generate Verilog Hardware Description Language (HDL) code for FPGA implementation [16]. It also provides Hardware/Software co-simulation within its environment to check validity of hardware design on actual device at initial design stage. Ethernet and JTAG is been supported to communicate between hardware and Simulink for performing Hardware/Software co-simulation. Thus XSG provides an ideal user friendly environment for modeling of emerging image processing algorithms on FPGA [10,11].

This paper focuses on an efficient implementation of local filters on FPGA by using XSG tool. Performance optimization of filter-design is achieved at two levels. At first level, Design space is explored on DSP friendly XSG environment to perform design-level optimization. It results in an efficient parallel and pipelined filter design that exploits tempo-spatial parallelism of local filters. Also computationally and memory intense portion of the design is efficiently implemented with DSP Slices and BRAM18 respectively. These powerful embedded FPGA resources [7,8] perform their dedicated functions efficiently for filter design. At second level, optimization efforts are extended to post-implementation level where efficient floor-planning fine-tunes the design. It aligns utilized embedded resources on target FPGA device i.e. Virtex5 and packs associated Slices across them. It results in a compact design with high level performance gain. The proposed two level optimized local filter design is suitable for potable, high-speed real time image processing applications.

Rest of the paper is organized as follows. Section 2 discusses related work. In Sect. 3 explains our local filter design. Section 4 declares results and comparison. Conclusion is included in Sect. 5.

2 Related Work

We have gone through a brief literature review on FPGA based local filters for image processing using XSG. In [5] a parallel architecture of local filters is designed for MRI image filtering by using Xilinx Block Sets. They implemented nine basic filters of size 5×5 by using Configurable Logic Blocks (CLBs). This architecture works at 228 MHz to process a 64×64 image on Virtex-6 FPGA. Also in [18] local filters (3×3 Sobel, 3×3 Prewitt and 2×2 Robert) were designed for edge detection by using Xilinx Block Sets and implemented them on FPGA by using CLBs for a 720×480 images. For a 3×3 Sobel filter, their design works at 154.9 MHz and occupies 954 Slices and 5 BRAMs on Spartan 3A DSP XCSD3400A FPGA. Similarly, [3] has also designed different local filters and declared results for edge enhancement. Their design occupies 436 Slices on Virtex-6 FPGA. Instead of using Xilinx Block Sets, [14] used Black Box to design 3×3 Sobel filter on XSG. The Black Box is used to incorporate HDL models into XSG. In their design, Black Box contains VHDL description for a 3×3 Sobel filter. They implemented Sobel filter for edge detection on Spartan 3 A and Virtex-5 FPGA by using CLBs. Their design occupies 117 Slices at the working frequency of 54.5 MHz on Spartan 3 A and 103.4 MHz on Virtex-5 FPGA.

Instead of implementing local filters on FPGA by using only CLBs (i.e. Slices), it is also possible to exploit powerful embedded resource of FPGA i.e. DSP Slices in combination with CLBs. DSP Slice is dedicated to efficiently perform computationally intense arithmetic operations at high operating frequency as compared to the conventional CLBs based logic. In [15] Sobel filter is implemented for video processing by using combination of DSP Slices and CLBs. It occupies 1284 Slices and 4 DSP Slices on Spartan 3A DSP FPGA. Their design works at 68.4 MHz. However, [19] implemented Sobel and Gaussian filters by using only DSP Slices on Virtex-4 FPGA. For 3×3 and 5×5 filters, their design occupies 11 and 26 DSP Slices respectively at the working frequency of 100 MHz.

All of the above discussed reported works [3,5,14,18] have implemented FPGA based local filters and performed design level optimizations. They implemented parallel architectures of local filters to accelerate their performance on FPGA. After design level optimization, [3,14,18] were not able to achieve significantly high operating frequencies for their design on low performance FPGA Spartan device. While [3,5,14,18] selected high performance FPGA device (Virtex5 and Virtex6) for implementing local filters with design level optimization but still not achieved significant performance gain. In [15], powerful dedicated DSP Slices were utilized to perform computationally intense multiplication operation of filter at a high operating frequency, but the rest of their design logic occupied large number of logic Slices. Consequently the inefficiently mapped bigger design with long routing delays limits overall design performance. In [19] DSP Slices were used for implementing overall filter design with the aim of increasing speed and accuracy of design but not shown significant performance gain.

The above discussed implementations attempted to accelerate performance of local filters on FPGA by performing design level optimizations but none of them has shown any optimization effort at post-implementation level that would

result in a compact local filter design with high performance gain. Therefor there is a need of an efficient design strategy which not only optimizes it at design level but also at post implementation level for portable and high speed image processing applications.

3 Our Work

In this work, we present an efficient implementation of local filters on FPGA by using XSG tool. A 3×3 Gaussian filter (coefficients are 1,2,1;2,4,2;1,2,1) and a 512×512 grey scale image of 8 bit precision has been chosen to explain our design concept. Figure 1 shows our design for a 3×3 Gaussian filter. It comprises of Row Buffers (RBs), Shift Registers (SRs), Coefficients ROM, Multiply and Accumulate unit (MACs), down sampler, adder tree, absolute conversion unit (abs) and convert unit (cvrt).

The RBs temporarily store consecutive input image rows to provide the MACs a seamless flow of pixels per clock cycle via SRs. The MACs multiplies these input pixels with pre-stored filter coefficients and accumulate them to yield partial results. These partial results are down sampled and summed up by using down sampler and adder tree respectively. Absolute value is computed by abs, and then final result narrows down to the required level by cvrt.

The proposed filter design performs two levels of optimization i.e. design level and post implementation level which results in a compact filter design with high degree of performance gain. The functionality of design is also validated by using Hardware/Software co-simulation method to increase design productivity. Detailed implementation of filter design along with two levels optimization is explained in following sub sections.

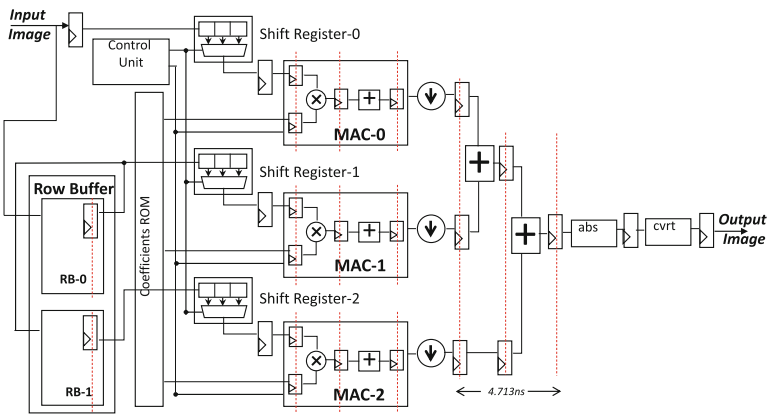


Fig. 1. Our proposed 3×3 filter design

3.1 Modeling of Local Filter Design Using XSG

We have used XSG for modeling of our filter design onto FPGA. The complete XSG model of our filter design is shown in Fig. 2. It comprises of four main units i.e. Input Unit, Internal Memory Unit, Arithmetic Unit and Output Unit. The Input Unit acquires a 512×512 input image and delivers pixels array to Internal Memory Unit. The Internal Memory Unit stores 2 consecutive rows of input image on FPGA by using 2 RBs. These RBs provide multiple pixels per clock cycle to Arithmetic Unit for parallel computation. The Arithmetic Unit performs convolution operation on input pixels and outputs filtered image pixels. The array of filtered pixels is converted back into 512×512 image matrix by using Output Unit. Detailed working of each unit is discussed separately in next sub-sections.

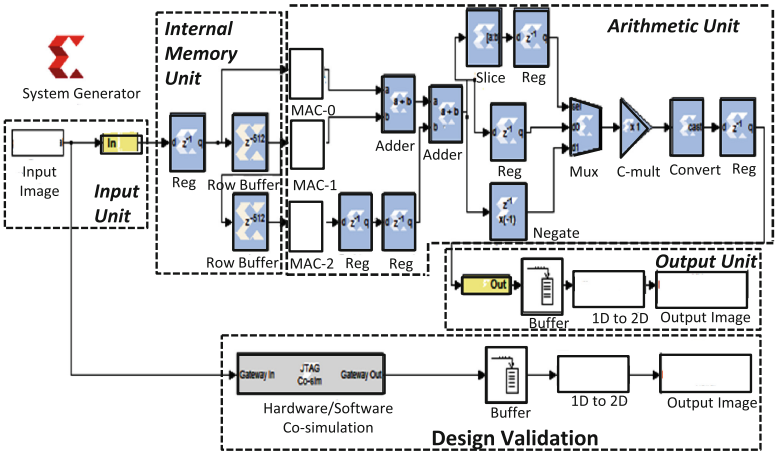


Fig. 2. Filter design modeled on Xilinx system generator

Input Unit. The Input Unit acquires a 512×512 grey scale image and behaviorally reshapes it from (row $\times$ column) matrix into (time $\times$ sample) matrix. The reshaped input image data has the time steps in first column and the pixels stream in second column. Since the XSG is a time based tool, thus the time step variable is implicitly considered by the input image. The Gateway In block of Input Unit converts double precision input image pixel data into a fixed point type and serves as an input port for the top level module of our HDL design.

Internal Memory Unit. Input image data via Input Unit is buffered in Internal Memory Unit. The 2 RBs in this unit are implemented with 2 BRAM18 primitives in simple dual port configuration. Each BRAM18 stores one complete row of 512×512 input image. With intake of single input pixel per clock cycle from external memory, the Internal Memory Unit provides multiple pixels per

clock cycle to parallel filter design for output pixels calculation at the throughput rate of 1clock/pixel.

Arithmetic Unit. Arithmetic unit is the main unit of our design which performs filtering operation (i.e. 2D convolution) on input image pixels. It is a parallel and pipelined architecture which exploits tempo-spatial parallelism of local filters. It comprises of three parallel connected 3tap MAC units, adders, absolute conversion unit (abs), convert unit and pipeline registers as shown in Fig. 2. Each MAC unit performs 1D convolution operation. It consists of DSP Slice, Op-mode, Addressable Shift Register (ASR), Distributed RAM (DRAM), Comparator, Counter, Capture Registers, Down Sampler and pipeline registers as shown in Fig. 3. The ASR implements SRs to shift input pixels over filter window (coefficients) for convolution operation. The filter coefficients are pre-stored in a DRAM implemented with FPGA Slices. A Comparator and a counter are required to implement address logic for ASR and DRAM. Counter generates addresses and Comparator creates the reset and enables signals for ASR and DRAM. The signal is asserted at 0 addresses and is delayed for pipeline stages. The DSP Slice is utilized to perform multiplication and accumulation operation upon getting instruction (P=P+(A*B)) by Op-mode. Output of DSP Slice is fed to a capture register that captures its final result after accumulating full set of filter samples. As sampling period of MAC is higher than desired output sampling period, so output from capture register is fed to a down sampler. It's down sampling rate is equal to the length of filter's coefficient array. For our case study i.e. a 3×3 filter, coefficient array length is 3 so the down sampling rate is also equal to 3.

Down sampled partial result from MAC units are then added to get a final output value. This is done by using a 2 stage adder tree structure. The absolute value of output is computed by using abs which is comprised of slice, negate and multiplexer (mux) blocks. Due to additions of partial results of parallel connected MAC units, the output value of adder tree is a larger bit number as compared to input. These results are rounded and quantized back to desired level. We used bit width convertor (convert) to narrow down final output to desired bit width. As we used 8 bit gray scale images at input, so our desired output level is set to 8

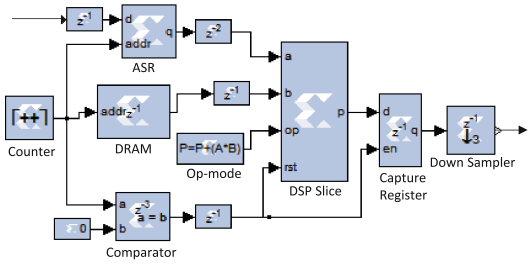


Fig. 3. MAC unit of our filter design

bit. The 8 bit filtered pixel of Arithmetic Unit in fixed point format is delivered to Output Unit of design.

Output Unit. The Output Unit acquires filtered output pixels from Arithmetic Unit. Its Gateway Out block converts fixed point output data to Simulink double precision value and serve as an output port for the top level module of the HDL design. The 1D output data stream then needs to convert back into 2D image matrix for display. A buffer is employed to convert scalar samples to frame at lower sampling rate, then conversion of 1D data into 2D image matrix is performed by using 1D to 2D conversion block in order to acquire output filtered image.

3.2 First Level Design Optimization

In first level of optimization, we explore design space of chosen local filter to exploit its inherent tempo-spatial parallelism. Firstly, for exploiting spatial parallelism, we design a parallel filter architecture in which the 2D filter operation (2D convolution) is partitioned into 3 parallel 1D filter operations (1D convolution). The 1D filter operation is performed simultaneously on input pixels of 3 rows to compute partial results. These partial results are then added up to the computation of final result.

As shown in Fig. 1, the input image pixels data from an external memory enters into RBs. For a 3×3 filter, 2 consecutive rows of input image are stored in 2 RBs implemented with 2 BRAM18 primitives (18 Kb BRAM). The use of RBs provide multiple pixels per clock cycle required by parallel filter design to exploit its spatial parallelism [9]. The RBs introduces an initial delay to our design which is equal to the time needed until the first 2 rows of the image are stored in RB. Subsequently, they provide required pixels per clock cycle to our filter design without any further delays. Pixels from RBs enter in SRs which are used to shift them over filter windows.

Input pixels via RBs and SRs, goes into 3 parallel connected 1D filters. The MAC units perform 1D filtering operation. Each MAC multiplies 3 pixels with 3 filter coefficients and accumulates them to give a single output pixel (i.e. partial result). Since MAC is the most computational intense unit of overall design, it is implemented with embedded DSP Slice which is a dedicated FPGA resource for performing high-speed arithmetic operations. The DSP Slice efficiently performs MAC operation at high operating frequency, low power consumption and high accuracy as compared to conventional Slice based MAC unit.

The output of MAC is sent to a down sampler unit (with down sampling rate equal to the length of filter's coefficient array i.e. 3 for said case study). Down sampled data of MAC units is accumulated by a (N−1) stage adder tree (i.e. $3 - 1 = 2$ for said case study). It accumulates all inputs and gives a filtered output pixel (i.e. final result). Absolute value is computed by abs, and then final result narrows down to required level by cvrt.

This parallel filter design bears a good performance aptitude however, the deployed adder tree structure incurs long critical path of 4.30 ns thus limiting

the overall design performance. To cater this bottleneck, we exploit temporal parallelism of adder tree structure in which its long combinational path is divided into smaller paths by inserting pipeline registers. These pipeline registers break long critical path associated with a combinational path of adder structure into small critical paths.

Critical paths associated with inputs and outputs of utilized embedded resources i.e. BRAM18 and DSP Slices in our design are also reduced by using their embedded primitive registers. These primitive registers hold the data of BRAM/DSP Slice in sub-pipeline manner to break longer critical path into shorter critical paths. We used output primitive registers of the BRAM and also multiplier primitive register of the DSP Slice. As these registers are embedded within these resources, they enhance overall performance of our design at no overhead hardware cost. By inserting pipeline and sub-pipeline stages in the design, its critical path is reduced to 2.48 ns thus increasing the overall performance (operating frequency) of design.

3.3 Second Level Design Optimization

The first level optimized design was modeled on XSG and then imported onto Xilinx ISE Environment. From here, we extend our optimization efforts for second level of optimization, in which we fine-tune the design by its device aware floor planning. Xilinx Synthesis Tool (XST) place and route our design on target Virtex5 device by using conventional Computer-Aided Design (CAD) mapping algorithms. These mapping algorithms are heuristic in nature that shows good results but not guarantee the optimal solution [12]. Therefore for optimal solution, besides using XST efficiently by choosing optimal mapping strategy for our design, we perform device aware manual floor planning of our design on target Xilinx Virtex5 device.

Our design is comprised of dedicated resources (BRAM18 and DSP Slice) in combination with CLBs (i.e. logic Slices). Both of the dedicated resources are available on target Virtex5 FPGA device in separate columns adjacent to each other as shown in Fig. 4. Also, pitch of BRAM tile (comprising two BRAM18 primitives) matches the DSP tile (comprising two DSP Slices i.e. DSP48E) [8]. We manually map both embedded resources of our design on FPGA such that the design creates a natural alignment due to alike tile pitches. Afterwards we pack the associated Slice based logic across utilized dedicated resources as shown in Fig. 5. This alignment of dedicated resources on device and placement of logic closely across these dedicated resources confined the overall design to a small device area. It not only results in further reducing critical path of confined design to 2.32 ns but also reduces occupied Slices of our design with reduced interconnect usage.

3.4 Design Validation

The functionality of our filter design is validated to improve its productivity for real time systems. We use Hardware/Software co-simulation to provide the

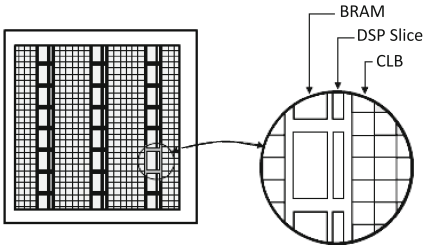


Fig. 4. Xilinx Virtex5 architecture

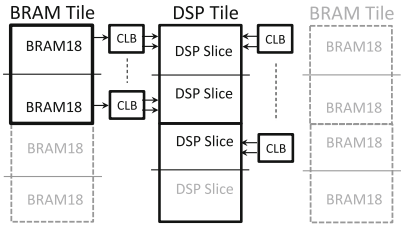


Fig. 5. Placement of our filter design on Virtex5

hardware version of design in flexible simulink environment. This hardware version interacts with the underlying FPGA device during simulation, performs necessary device configuration, data transfer and clocking. The JTAG cable is used to communicate between FPGA device and Simulink. The bit stream generated is verified by downloading it onto the FPGA board and ensuring that it runs successfully.

4 Results and Discussion

This work efficiently implements local filters on FPGA device using XSG. A 512×512 grey scale image and a 3×3 Gaussian filter were considered as a case study. The design was tested and evaluated on Virtex-5 (Xc5vfx100tff1738-3) FPGA device within the ISE 13.1 development Suite. The test case filtration on a standard Lena image by using our design is shown in Fig. 6.

Table 1 shows implementation results of our filter design in terms of device utilization and operating frequency. The design after first levels optimization utilizes 2 BRAM18, 3 DSP Slices and 145 Slices at operating frequency of 411.18 MHz. The results are improved by performing second level optimization. After second level optimization, Slice count of design is reduced by 38.6 % with a 5 % performance gain. The design after two level of optimization fits in just 89 Slices, 2 BRAM18 and 3 DSP Slices and achieves high operating frequency of 431.03 MHz. Since the filter is implemented in a parallel and pipelined way, after an initial latency it produces an output at every clock cycle. Therefore the design at this frequency is capable to sustain a throughput rate of 431.03 Mega Operations

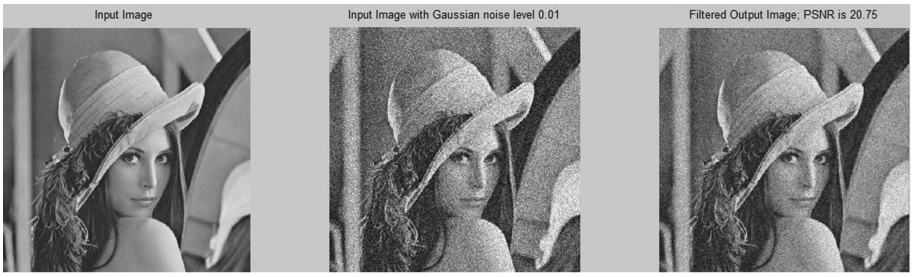


Fig. 6. Test case Gaussian Filtering on a standard Lena input image

Table 1. Implementation results for said case study on Virtex5 (Xc5vfx100tff1738-3).

FPGA resources	First level optimization	Second level optimization	% Improvement
Slices	145	89	38.6 %
DSP Slice	3	3	0
BRAM18	2	2	0
Frequency (MHz)	411.18	431.13	5 %

per Second (MOS). With this throughput, for high definition image resolution 1920×1080, the architecture performs at 208 frames per second far exceeds the real time image processing requirements.

Comparison of our results with other reported results [15,18,19] is shown in Table 2. Since these results were available on different FPGA devices, therefore for a fair comparison we provide results in terms of equivalent slice count and equivalent frequency which is calculated by multiplying reported results with a normalizing factor. Firstly, the normalizing factor for slice count is calculated on basis of the equivalent logic which can be synthesized by the logic resources of different Xilinx FPGA devices. One Slice of Virtex-4 and below families is comprised of two LUT4 to perform 32 lookup operations while One Slice of Virtex-5 and above families is comprised of four LUT-6 to perform 256 lookup operations therefore a factor of 8 (256/32) is applied as normalizing factor for their comparison [2] i.e. 1 Slice of Virtex-5 families or above is equivalent to 8 Slices of Virtex-4 or below families. Secondly the normalizing factor for frequency is also calculated by taking ratio of maximum achievable frequency of two dissimilar FPGA devices under comparison. The maximum achievable frequency of these designs on Spartan 3, Virtex-4, Virtex-5, and Virtex-6 are 320 MHz, 500 MHz, 550 MHz and 600 MHz therefore the frequency normalizing factors with respect to target Virtex-5 device are 1.71, 1.1 and 0.9 respectively. Moreover, besides considering the equivalent slice count and equivalent frequency for eliminating the impact of technology difference on results, the size of all reported designs are also scaled to a 3×3 size for eliminating difference in results due to different design sizes. The results of scaled design in terms of equivalent Slice count

Table 2. Results comparison.

Work	Device	Equivalent frequency (MHz)	Equivalent slices	DSP slices
[15]	Spartan3	117.5	81	4
[18]	Spartan3A DSP	266.2	60	b
[19]	Virtex4	110	a	11
[5]	Virtex6	210.8	a	b
Proposed	Virtex5	**431.13**	**89**	3

[a]Not Given; [b]Not Used

and equivalent frequency is compared in Table 2. It is evident that our design outperforms all other implemented designs in terms of operating frequency and at the same time occupies reduced FPGA resources.

5 Conclusion

This paper presented an efficient implementation of local filters on FPGA. Performance optimization of filter-design is achieved at two levels. Firstly at the design level by exploiting tempo-spatial parallelism and secondly at post-implementation level by efficient device aware floor-planning which further fine-tunes the design and confine the design across dedicated embedded resources. The proposed two level optimized local filter design with low resource consumption and high operating frequency is a suitable solution for portable, high-speed ($1920 \times 1080@208$) real time applications in the field of computer vision, medicine, military and industry to name a few.

References

1. Bailey, D.G.: Design for Embedded Image Processing on FPGAs. John Wiley & Sons, Singapore (2011)
2. Betz, V., Rose, J., Marquardt, A.: Architecture and CAD for Deep-Submicron FPGAs, vol. 497. Springer Science & Business Media, New York (2012)
3. Elamaran, V., Praveen, A., Reddy, M.S., Aditya, L.V., Suman, K.: FPGA implementation of spatial image filters using Xilinx system generator. Proc. Eng. **38**, 2244–2249 (2012)
4. Gonzalez, R.C., Woods, R.E.: Digital Image Processing. Prentice Hall, Upper Saddle River (2002)
5. Hasan, S., Yakovlev, A., Boussakta, S.: Performance efficient FPGA implementation of parallel 2-D MRI image filtering algorithms using Xilinx system generator. In: 2010 7th International Symposium on Communication Systems Networks and Digital Signal Processing (CSNDSP), pp. 765–769. IEEE (2010)
6. Hedberg, H.: Image processing architectures for binary morphology and labeling. Lund University (2008)
7. Xilinx Inc.: Virtex-5 FPGA xtremedsp design considerations user guide, January 2009

8. Xilinx Inc.: Virtex-5 FPGA User Guide v5.4 (2012)
9. Kazmi, M., Aziz, A., Akhtar, P., Kundi, D.E.S.: FPGA based compact and efficient full image buffering for neighborhood operations. Adv. Electr. Comput. Eng. **15**(1), 95–104 (2015)
10. Kiran, M., War, K.M., Kuan, L.M., Meng, L.K., Kin, L.W.: Implementing image processing algorithms using hardware in the loop approach for Xilinx FPGA. In: International Conference on Electronic Design, ICED 2008, pp. 1–6. IEEE (2008)
11. Moreo, A.T., Lorente, P.N., Valles, F.S., Muro, J.S., Andrs, C.F.: Experiences on developing computer vision hardware algorithms using Xilinx system generator. Microprocess. Microsyst. **29**(8), 411–419 (2005)
12. Oklobdzija, V.G.: The Computer Engineering Handbook. CRC Press, Boca Raton (2001)
13. Perri, S., Lanuzza, M., Corsonello, P., Cocorullo, G.: A high-performance fully reconfigurable FPGA-based 2D convolution processor. Microprocess. Microsyst. **29**(8), 381–391 (2005)
14. Said, Y., Saidani, T., Smach, F., Atri, M.: Real time hardware co-simulation of edge detection for video processing system. In: 2012 16th IEEE Mediterranean Electrotechnical Conference (MELECON), pp. 852–855. IEEE (2012)
15. Said, Y., Saidani, T., Smach, F., Atri, M., Snoussi, H.: Embedded real-time video processing system on FPGA. In: Elmoataz, A., Mammass, D., Lezoray, O., Nouboud, F., Aboutajdine, D. (eds.) ICISP 2012. LNCS, vol. 7340, pp. 85–92. Springer, Heidelberg (2012). doi:10.1007/978-3-642-31254-0_10
16. Saidani, T., Atri, M., Dia, D., Tourki, R.: Using Xilinx system generator for real time hardware co-simulation of video processing system. In: Ao, S.-L., Gelman, L. (eds.) Electronic Engineering and Computing Technology. LNEE, vol. 60, pp. 227–236. Springer, Netherlands (2010)
17. Samarawickrama, M.G.: Performance Evaluation of Vision Algorithms on FPGA. Universal-Publishers (2010)
18. Sudeep, K., Majumdar, J.: A novel architecture for real time implementation of edge detectors on FPGA. Int. J. Comput. Sci. Issues **8**(1), 193–202 (2011)
19. Wasfy, W., Zheng, H.: General structure design for fast image processing algorithms based upon FPGA DSP slice. Phys. Proc. **33**, 690–697 (2012)
20. Woods, R., McAllister, J., Lightbody, G., Yi, Y.: FPGA-Based Implementation of Signal Processing Systems. Wiley Online Library (2008)

A Frequency Dependent Investigation of Complex Shear Modulus Estimation

Quang Hai Luong[1(✉)], Manh Cuong Nguyen[1], and Tran Duc Tan[2]

[1] Biomedical Engineering Department, Le Quy Don Technical University, 236 Hoang Quoc Viet Street, Cau Giay District, Hanoi, Vietnam
luonghai@mta.edu.vn
[2] VNU University of Engineering and Technology, 144 Xuan Thuy Street, Cau Giay District, Hanoi, Vietnam
http://www.mta.edu.vn

Abstract. Mechanical properties of tissues in terms of elasticity and viscosity provide us useful information which may be used in detecting tumors. Shear wave imaging (SWI) is a new method to quantify tissue elasticity by estimating the parameters of the complex shear modulus (CSM). The shear wave is generated by a vibrating needle at a certain frequency. In fact, CSM is a function of the vibrating frequency. Therefore, in this paper, a frequency dependent investigation of CSM will be carried in order to evaluate the estimation performance. The Extended Kalman Filter (EKF) is designed to estimate the CSM for both homogeneous and heterogeneous mediums. The root mean square (rms) error is used to evaluate the quality of the CSM estimation. Several tests were implemented to determine the range of vibrating frequency should be used for the good estimation.

Keywords: Shear wave elasticity imaging · Complex shear modulus · Extended Kalman Filter · Vibration

1 Introduction

Mechanical properties of tissues in terms of elasticity and viscosity provide us useful information which may be used in medical diagnosis, especially in detecting tumors [1]. Among various elasticity imaging modalities, ultrasonic shear wave elasticity imaging (SWEI), introduced in 1998 by Sarvazyan et al. [2], is used for estimating the complex shear modulus (CSM) of biphasic hydro polymers including soft biological tissues. As a consequence, SWEI can be coupled with traditional (e.g., structural) ultrasound imaging to provide additional information in the diagnosis. In a recent survey on different state-of-art techniques of ultrasound elastography [3], Gennission et al. have confirmed that SWEI has significant advantages over the other techniques in terms of reproducibility, quantification, elasticity contrast, and automatic shear wave generation. These advantages lead to new applications of SWEI, not only for diagnosis but also for treatment.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_6

With respect to CSM estimation, various methods have been developed as briefly surveyed next. In 2004, by using the fact that propagation speed of shear waves is related to the frequency of vibration, the elasticity and viscosity of the medium Chen et al. proposed a method to estimate the shear elasticity and viscosity of a homogeneous medium by measuring the shear wave speed dispersion and, in turns, the CSM [4]. In 2007, Zheng et al. applied a linear Kalman filter for the reconstruction of the harmonic motion of particle velocities at distinct spatial locations. Their approach is to model displacement at the spatial points of interest as a sinusoidal function of time. From estimated quantities, absolute phase at a distinct spatial location can be found. By repeating the same procedure for another location a phase difference is found. Shear wave speed and shear wave dispersion curves are estimated over a frequency bandwidth and material properties are obtained. The stochastic filtering approach helped the authors to obtain optimal estimates of the temporal phase at the given spatial location. A drawback of this method is that, the CSM reconstruction is not optimal and is a post-processing procedure requiring several shear wave frequency measurements [5]. In 2010, Orescanin et al. have conducted an experiment whereby they modeled the nonlinear relationship between wave dynamics and material parameters. They represented the CSM parameters of the present by a nonlinear function of the CSM parameters in the past. So, they applied the Maximum Likelihood Ensemble Filter (MLEF), which is a stochastic filter capable of handling nonlinear dynamical models and nonlinear observation operators, to estimate the CSM of a homogeneous medium based on the Kelvin–Voigt model [6]. In this study, they investigated the change of the wave number and attenuation of the wave propagation of shear waves when the frequency and amplitude of vibration were changed. This approach has been extented to a heterogeneous medium [8].

Currently, for the problem of the CSM estimation, there are two key tasks. First, estimating accurately both the elasticity and the viscosity. Second, the CSM estimation can be performed in an online manner during data acquisition. In this paper, we applied the EKF to estimate the CSM. Moreover, we investigated the impact of the frequency of vibration on the quality of the CSM estimation for different type of soft tissues.

2 The Methods

2.1 Shear Wave Propagation and Generation

Shear wave is generated and measured as according to Fig. 1. A mechanical actuator was adapted to hold a stainless-steel needle. The needle is 1.5 mm in diameter and 13 cm long. It is controlled to vibrate along the z axis with frequency from 100 Hz to 500 Hz. So the shear wave is propagated in tissue. The Doppler ultrasound system was used to measure the particle velocity [6]. The needle vibrates along the vertical (z) axis. Under an assumption of cylindrical shear wave propagation along the radial axis, the particle velocity of shear wave

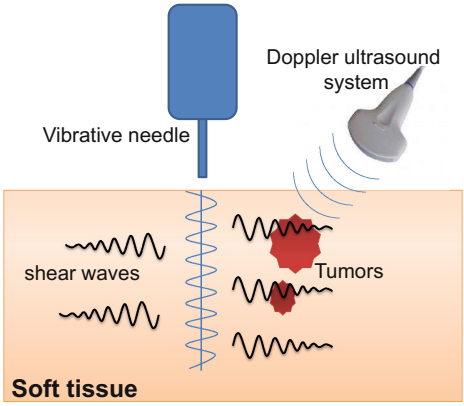


Fig. 1. Generation and measurement shear wave.

$v(r,t)$ is a spatial-temporal function of the radial distance r and time t, and is given by

$$v(r,t) = \frac{1}{\sqrt{r-r_0}} A e^{-\alpha(r-r_0)} \cos[\omega t - k_s(r-r_0) - \phi], \tag{1}$$

where A is the amplitude of source excitation, r_0 is the needle position, and ϕ is the initial temporal phase, α and k_s are attenuation coefficient and wave number at surveyed point. The particle velocities at every spatial location is measured by using the Doppler acquisition.

2.2 The Impact of Frequency of the Vibration on the Shear Wave Propagation

According to Eq. (1), the attenuation coefficient α has effect significantly on the shear wave propagation. The more α is great, the more the particle velocity of shear wave is attenuated. Essentially, α and k_s are imaginary and real components of complex wave number k_s'.

$$k_s' = k_s + i\alpha. \tag{2}$$

On the other hand, follow Kelvin–Voigt model, we have

$$c_s = \sqrt{\frac{\mu}{\rho}}, \tag{3}$$

$$\mu = \mu_1 - i\omega\eta, \tag{4}$$

where c_s is shear wave speed, ρ is mass density of medium (tissue) at the surveyed point, μ is the viscoelasticity, μ_1 and η are the elasticity and viscosity of medium at the surveyed point. Complex wave number k_s' is defined as:

$$k_s' = \frac{2\pi f}{c_s}. \tag{5}$$

From Eqs. (2), (3), (4) and (5), and replace $\omega = 2\pi f$, we have

$$k_s' = k_s + i\alpha = 2\pi f \sqrt{\frac{\rho}{\mu_1 - i(2\pi f)\eta}}. \tag{6}$$

According to Eq. (6), for every tissues (this means that the valued ranges of the elasticity μ_1 and viscosity η of medium are determined), the change of the frequency of vibration lead to the change of the attenuation coefficient α, in turns, the shear wave propagation.

2.3 Estimating the CSM Using the EKF

The CSM estimation is synonymous with estimating the elasticity μ_1 and the viscosity η. From Eq. (6), μ_1 and η at each point are calculated follow formulas

$$\mu_1 = \frac{\rho\omega^2(k_s^2 - \alpha^2)}{(k_s^2 + \alpha^2)^2}, \tag{7}$$

$$\eta = \frac{2\rho\omega k_s \alpha}{(k_s^2 + \alpha^2)^2}. \tag{8}$$

So the problem of the CSM estimation becomes one of the wave number and attenuation coefficient estimation. In this paper, we applied the EKF to estimate the wave number and attenuation coefficient. At each point in space, we built a system model which is used in the Extended Kalman problem. In discrete form, Eq. (1) becomes

$$v_n(r) = \frac{1}{\sqrt{r - r_0}} A e^{-\alpha(r - r_0)} \cos[\omega n \Delta t - k_s(r - r_0) - \phi], \tag{9}$$

where n is the discrete time index and Δt is sampling cycle. Transforming trigonometric Eq. (9), we receive

$$\begin{aligned} v_n(r) = v_{n-1}(r)\cos(\omega\Delta t) - \tfrac{1}{\sqrt{r - r_0}} A e^{-\alpha(r - r_0)} \\ \sin[\omega(n-1)\Delta t - k_s(r - r_0) - \phi]\sin(\omega\Delta t). \end{aligned} \tag{10}$$

To estimate the attenuation coefficient α and wave number k_s using the EKF, Eq. (10) is written in state equation form

$$x_n = f(x_{n-1}, p_{n-1}). \tag{11}$$

Equation (11) is equivalent to

$$\begin{bmatrix} v_n \\ \alpha_n \\ k_{s(n)} \end{bmatrix} = \begin{bmatrix} F(v_{n-1}) \\ \alpha_{n-1} \\ k_{s(n-1)} \end{bmatrix}, \tag{12}$$

where $x_n = \begin{bmatrix} v_n \\ \alpha_n \\ k_{s(n)} \end{bmatrix}$ is state vector at each point, the random variable p_n is process noise, F is a non-linear function, F describes the relation between v_{n-1}

and v_n as shown in Eq. (10), $\alpha_{n-1} = \alpha_n$ and $k_{s(n-1)} = k_{s(n)}$ because we assume that α and k_s would not be changed during the time of the experiment. By using the Doppler acquisition, the measured particle velocities at every spatial location are impacted by Gaussian noise $w_n(r)$. So the measured particle velocity is

$$\widehat{v}_n = v_n + w_n. \tag{13}$$

To use the EKF, Eq. (13) is written in measurement equation form of Kalman problem

$$y_n = h(x_n, w_n). \tag{14}$$

Equation (14) is equivalent to

$$\hat{v}_n = \begin{bmatrix} 1 & 0 & 0 \end{bmatrix} \begin{bmatrix} v_n \\ \alpha_n \\ k_{s(n)} \end{bmatrix} + w_n, \tag{15}$$

where $y_n = \widehat{v}_n$ is measurement vector at each point. From Eqs. (11) and (14), x_n is estimated by using the EKF according to the algorithm in [7]. The result is that the shear wave attenuation coefficient α and the wave number k_s at each point are estimated.

3 Results and Discussions

We built some simulation scenarios to test the proposed methods. First, we created three type of soft tissues T1, T2 and T3 (like liver, breast, prostate). Their CSM (μ_1 and η) are (2000 Pa and 0.2 Pa.s), (18000 Pa and 0.5 Pa.s) and (36000 Pa and 1 Pa.s), respectively. The above values, we refer to data table which was shown in [1]. For every type of tissues, we investigated the impact of the frequency of vibration on the attenuation coefficient follow Eq. (6). The results are shown in Fig. 2, which indicates that the attenuation coefficient for liver tissue, breast tissue and prostate tissue are nearly equal (approximates 20) at frequencies 200 Hz, 400 Hz and 450 Hz, respectively. Next, we created three heterogeneous mediums which simulate three above tissues. The details of the mediums and tumors are shown in Table 1. The surveyed number of points are 43 (they are on a line, the distance between 2 points is 0.3 mm), the coordinate of the vibration needle is 0 mm, $r_0 = 0.3$ mm, the amplitude at zero location $A = 2$ mm, the mass density of medium $\rho = 1000\,\text{kg/m}^3$, $\Delta t = 0.06$ ms, time step's number $n = 500$. the vibration frequency of the needle is changed from 100 Hz to 500 Hz (step 50 Hz).

We estimated the CSM using the EKF for all three above medium. For T1 tissue, Fig. 3 shows the estimated particle velocity and the noised particle velocity in time. The estimated particle velocity in time was filtered effectively. It is as a sinusoidal function of time.

The estimated particle velocity in space is indicated in Fig. 4. It is an attenuated-sinusoidal function and similar to the ideal particle velocity,

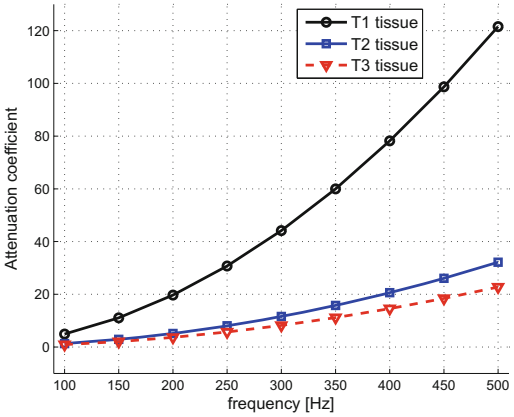


Fig. 2. For some type of tissues, the relation between the attenuation coefficient α and the frequency of vibration f.

Table 1. Parameters of tumors

Type of tissue	μ_1 (Pa)	η (Pa.s)	Location of tumor	μ_1 of tumor (Pa)	η of tumor (Pa.s)
T1	4000	0.2	(10 to 20)	5000	0.4
T2	18000	0.5	(10 to 20)	20000	0.6
T3	36000	1	(10 to 20)	42000	1.1

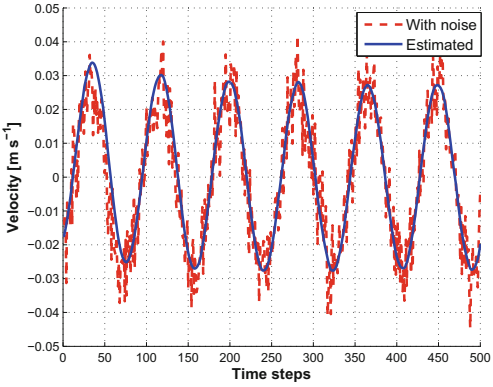


Fig. 3. Estimated particle velocity in time, with SNR = 20 dB, for T1 tissue.

while the noised particle velocity goes up and down in every short segment. This can confirm that the particle velocity in space was estimated effectively.

Fig. 5 indicates the estimated wave number in space, while Fig. 6 shows the estimated attenuation coefficient in space. The results are very good. Visually, the wave number is estimated better than the attenuation coefficient.

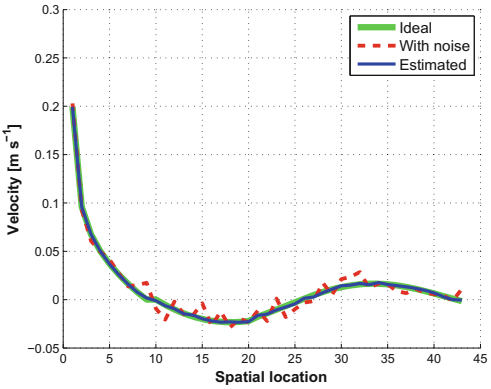


Fig. 4. Estimated particle velocity in space, with SNR = 20 dB, for T1 tissue.

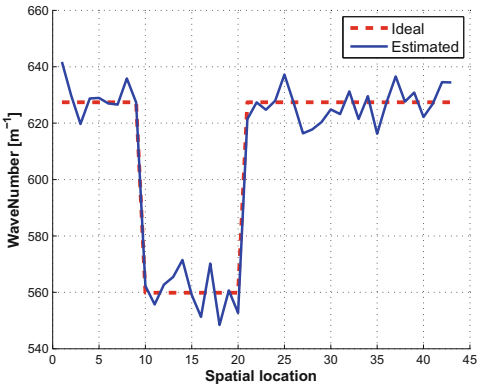


Fig. 5. Estimated wave number in space, for T1 tissue.

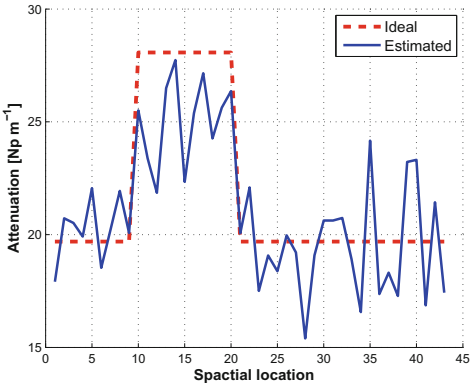


Fig. 6. Estimated attenuation coefficient in space, for T1 tissue.

To evaluate the quality of the CSM estimation when the frequency of vibration is changed, we used the *rms* error. The *rms* error is defined as in Eq. (16).

$$rms = \sqrt{\frac{\sum_{i=1}^{N} (\widehat{x}_i - x_i)^2}{N}}, \tag{16}$$

where x is ideal input, $\widehat{x}$ is the estimated value, N is the number of samples. For T1 tissue, Fig. 7 shows the change of the rms error for the estimated wave number and attenuation coefficient following the frequency of vibration. At frequency 200 Hz, the values of the rms error for both the wave number and attenuation coefficient are least. This means that the quality of the CSM estimation is best.

Of course, we tried to build the parameters of the EKF, with which, the quality of the CSM estimation is best at 200 Hz for T1 tissue. However, we applied this EKF to estimate the CSM for T2 and T3 tissues, the results illustrated that

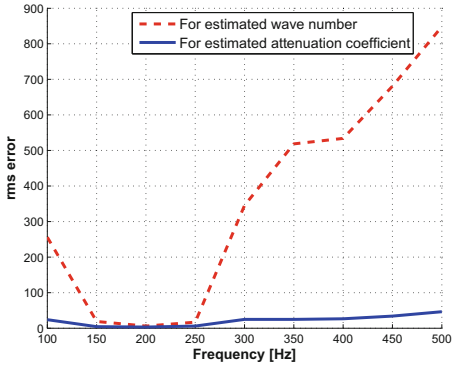


Fig. 7. The *rms* error for estimated wave number and attenuation coefficient, for T1 tissue.

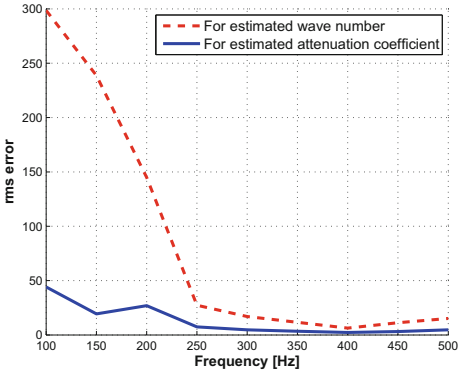


Fig. 8. The *rms* error for estimated wave number and attenuation coefficient, for T2 tissue.

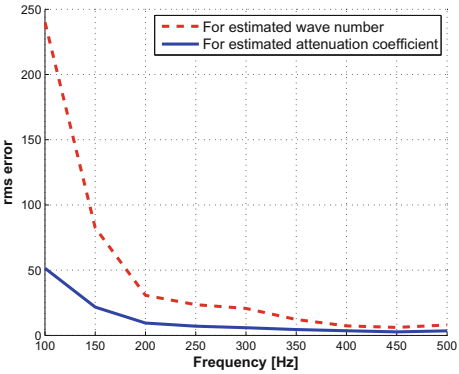


Fig. 9. The *rms* error for estimated wave number and attenuation coefficient, for T3 tissue.

the quality of the CSM estimation is best at 400 Hz (Fig. 8) and 450 Hz (Fig. 9) for T2 and T3 tissues, respectively.

4 Conclusions

This paper was successful in investigating the frequency dependent on estimating CSM for a heterogeneous medium. An effective EKF is designed to estimate the CSM at each point of a line in tissues. By extending tens of points in a line, we could estimate the CSM for a one-dimensional heterogeneous medium. Based on simulated results, we can determine the range of vibrating frequency should be used for the good estimation. In the future work, we will expand to a two-dimensional heterogeneous medium (i.e. CSM imaging). The accuracy of the CSM estimation would also be concerned, especially the viscosity of tissues.

References

1. Bercoff, J., Criton, A., Bacrie, C., Souquet, J., Tanter, M., Gennisson, J., Deffieux, T., Fink, M., Juhan, V., Colavolpe, A. et al.: ShearWave elastography: a new real time imaging mode for assessing quantitatively soft tissue viscoelasticity. In: Ultrasonics Symposium, IUS IEEE, pp. 321–324 (2008)
2. Sarvazyan, A.P., Rudenko, O.V., Swanson, S.D., Fowlkes, J.B., Emelianov, S.Y.: Shear wave elasticity imaging: a new ultrasonic technology of medical diagnostics. Ultrasound Med. Biol. **24**(9), 1419–1435 (1998)
3. Gennisson, J.L., Deffieux, T., Fink, M., Tanter, M.: Ultrasound elastography: principles and techniques. Diagn. Interv. Imaging **94**(5), 487–495 (2013)
4. Chen, S., Fatemi, M., Greenleaf, J.F.: Quantifying elasticity and viscosity from measurement of shear wave speed dispersion. J. Acoust. Soc. Am. **115**(6), 2781–2785 (2004)

5. Zheng, Y., Chen, S., Tan, W., Kinnick, R., Greenleaf, J.: Detection of tissue harmonic motion induced by ultrasonic radiation force using pulse-echo ultrasound and Kalman filter. IEEE Trans. Ultrasonics Ferroelectr. Freq. Control **54**(2), 290–300 (2007)
6. Orescanin, M., Insana, M.F.: Model-based complex shear modulus reconstruction: a Bayesian approach. In: IEEE Ultrasonics Symposium (IUS), pp. 61–64. IEEE Press (2010)
7. Welch, G., Bishop, G.: An Introduction to the Kalman filter. University of North Carolina, Chapel Hill (2006)
8. Tran-Duc, T., Wang, Y., Linh-Trung, N., Do, M.N., Insana, M.F.: Complex shear modulus estimation using maximum likelihood ensemble filters. In: 4th International Conference on Biomedical Engineering in Vietnam, pp. 313–316. Springer, Berlin (2013)

A Method to Enhance the Remote Sensing Images Based on the Local Approach Using KMeans Algorithm

Trung Nguyen Tu(✉), Duc Dang Van, Huy Ngo Hoang, and Thoa Vu Van

Institute of Information Technology, Vietnam Academy of Science and Technology, Post and Telecommunications Institute of Technology, Hanoi, Vietnam
{nttrung,dvduc,nhhuy}@ioit.ac.vn, thoa236@gmail.com
http://ioit.ac.vn

Abstract. The image enhancement methods based on fuzzy logic make image which quality higher clearly the traditional methods. However, actually, the methods still use the global approach, so having difficulty to enhance all land covers in remote sensing images. This paper presents a local approach based new algorithm of image enhancement for the remote sensing images, calculating thresholds automatically and combination multiple gray adjust operators.

Keywords: Image enhancement · Remote sensing images · Local approach · KMeans

1 Introduction

Remote sensing images often have large size and high resolution. However, they can also contain noises. To reduce noises and enhance the image quality, some methods should be used. Quality Enhancement, a necessary step in image processing, involves two separate phases: Image Enhancement and Restore. Remote sensing images include common noises and specific noises. To reduce specific noises composed of mist and cloud, we need specific methods by using Mallat algorithm [4]. In term of common noise reduction, common methods such as noise filter, image smoothing; contrast enhancement, adjusting gray levels of images should be adopted. Many common contrast enhancement methods apply the global approach to enhance all brightness levels of images. However, these methods are difficult to enhance all land covers in remote sensing images because they may lose the local contrast information and details in bright or dark regions. In [2,8], the authors combine the fuzzy logic [1] and the gray level adjusting formulas to enhance contrast of medical images. The methods consider membership matrix and the gray level adjusting formulas to enhance contrast. Yet the methods use global approach, so they cannot solve the problem of traditional methods. In addition, the fuzzy image enhancing methods still select values of the upper, lower, and mean thresholds manually. Therefore, if selecting the bad values, which may negatively affect the result of image enhancing.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_7

In this study, we propose a new algorithm of the image enhancement based on the local approach which combines the auto thresholds computing following each cluster, a gray level adjust model with using KMeans clustering algorithm.

The remaining sections of this report are presented as follow: Sect. 2 presents the algorithm of image enhancement based on fuzzy logic. Section 3 shows the algorithm of remote sensing enhancement based on global approach. Experiments and accessments are presented in Sect. 4. Section 5 is the conclusion of the article.

2 Image Enhancement Based on Fuzzy Logic

2.1 Overview of Image Enhancement Based on Fuzzy Logic

Fuzzy image enhancement is based on gray level mapping into a fuzzy plane, using a membership transformation function [7,8]. The aim is to generate an image of higher contrast than the original image by giving a larger weight to the gray levels that are closer to the mean gray level of the image than to those that are farther from the mean [2]. In recent years, many researchers have applied the fuzzy set theory to develop new techniques for contrast improvement [2]. An image I with its size of M x N and L gray levels can be considered as an array of fuzzy singletons, each having a value of membership denoting its degree of brightness relative to some brightness levels. For an image I, we can write in the notation of fuzzy sets:

$$I = U_{mn}\mu_{mn}/g_{mn} \tag{1}$$

where g_{mn} is the intensity of (m, n)th pixel and μ_{mn} its membership value. The membership function characterizes a suitable property of image (e.g. edginess, darkness, textural property) and can be defined globally for the whole image or locally for its segments. In recent years, some researchers have applied the concept of fuzziness to develop new algorithms for image enhancement. The principle of fuzzy enhancement scheme is illustrated in Fig. 1.



Fig. 1. The Main Principles of FuzzyImage Enhancement [2].

In the Subsects. 2.2 and 2.3, the article presents the algorithm of fuzzy image enhancement by using intensification operators and Hyperbol.

2.2 Fuzzy Image Enhancement with Intensification

This method uses the intensification operator [9] to reduce the fuzziness of the image which results in an increase of image contrast [10,11]. The algorithm is described as follows:

Step 1: Setting the parameters F_e, F_d, g_{max} of membership function

$$F_e = 2 \tag{2}$$

$$F_d = \frac{g_{max} - g_{mn}}{0.5^{-1/F_e} - 1} \tag{3}$$

Step 2: Define the membership function

$$\mu_{mn} = G(g_{mn}) = [1 + \frac{g_{max} - g_{mn}}{F_d}]^{-F_e} \tag{4}$$

Step 3: Modify the membership values

$$\begin{cases} 2[\mu_{mn}]^2 & 0 \leq \mu_{mn} \leq 0.5 \\ 1 - 2[1 - \mu_{mn}]^2 & 0.5 \leq \mu_{mn} \leq 1 \end{cases} \tag{5}$$

Step 4: Generatenew gray-levels

$$g'_{mn} = G^{-1}\left(\mu'_{mn}\right) = g_{mn} - F_d\left(\left(\mu'_{mn}\right)^{\frac{-1}{F_e}} - 1\right) \tag{6}$$

2.3 Fuzzy Image Enhancement with Hypebol Operator

The idea of histogram hyperbolization, and fuzzy histogram hyperbolization is described in [12] and [13], respectively. Due to the nonlinear human brightness perception, this algorithm modifies the membership values of gray levels by a logarith function. The algorithm can be formulated as follows:

Step 1: Setting the shape of membership function.

Step 2: Setting the value of the fuzzifier β.

Step 3: Calculation of membership values μ_{mn}.

Step 4: Modification of the membership values by β.

Step 5: Generation of new gray levels, as described below.

The choice of the membership function is very important, as the membership function characterize a certain property of the image (edginess, darkness, textual property). In this algorithm the shape of membership function is set as a triangular to characterize the hedges, and the value of fuzzifier β as a linguistic

hedge such that: $\beta = -0.75 + \mu 1.5$. Then by calculating the membership values μ_{mn} and modifying the membership values by β. Generate new gray levels values g_{mn} by following equation:

$$g'_{mn} = \left(\frac{(L-1)}{e^{-1}-1}\right)\left(e^{-\mu_{mn}-g_{mn}^{\beta}-1}\right) \tag{7}$$

2.4 Disadvantages of the Image Enhancement Algorithms Based on Fuzzy Logic

With a description of the fuzzy image enhancement method as above, we have the following comments:

- The first, the fuzzy image enhancement algorithms still use global approach as the traditional methods, advantages of this method is only adding a membership matrix to image. Therefore, this method is still difficult to enhance all land covers in remote sensing images because the local contrast information and details can still be lost in bright and dark regions.
- The second, when implementing algorithm, for example in [2], the thresholds of upper bounds max, lower bounds min still must be setted manually with each image. So, it is not effective.
- The third, like the traditional methods, the fuzzy image enhancement methods only performs with a spectral channel. With multispectral images as color images RGB or remote sensing images, algorithm is performed on each channel. Suppose, having a pixel P with gray levels correspond to the channels as (g_1^P, g_2^P, g_3^P ...). So, gray values g_1^P, g_2^P, g_3^P ... have relation of the same pixel. When performing fuzzy image enhancement algorithm follow each channel, this relation is not covered. Therefore, each gray value in set (g_1^P, g_2^P, g_3^P ...) is enhanced independently so it is hard to ensure the above spectral relation of new gray values after enhancing. Therefore, color of objects in output image can not be conserved.

3 The Remote Sensing Image Enhancement Based on the Local Approach

To overcome the disadvantages in Subsect. 3.4, we propose a new remote sensing enhancement method based on the local approach according to cluster. Accordingly, gray levels are enhanced to follow each cluster. The algorithm of remote sensing image enhancement based on local approach is presented in Subsect. 4.1.

3.1 Local Based Remote Sensing Image Enhancement

The main steps of algorithm of Local based Remote Sensing Image Enhancement using KMeans algorithm (LoRSIEK) are listed in Table 1.

Table 1. Processing of LoRSIEK.

Step	Task
1	Localizating remote sensing image based on clustering algorithm KMeans
2	Constructing model of adjusting gray level in accordance with cluster
3	Calculating automatic thresholds in accordance with each cluster
4	Generating enhanced image based on the gray level adjusting in accordance with each cluster

a-Localizating Remote Sensing Image Based on Clustering Algorithm KMeans. Input: n object and number of clusters k. Output: Clusters $V_i (i = 1..k)$ so that the follow objective function E reachs minimum:

$$E = \sum_{i=1}^{k} \sum_{x \in V_i} d^2 (x, m_i) \tag{8}$$

KMeans algorithm [18] includes 4 steps:

Step 1: Initialization.

Step 2: Assigning cluster center according to distance with (9).

$$d (x, V_S) = mind (x, V_j), 1 \leq j \leq k \tag{9}$$

Step 3: Updating the cluster center with (10).

$$V_j = \frac{\sum_{x \in Cluster_j} x}{count (Cluster_j)} \tag{10}$$

Step 4: Repeating and testing condition to stop.

b-Constructing Model of Adjusting Gray Level in Accordance with Cluster. In this subsection, we will construct function of gray level adjusting to enhance in accordance with each cluster. This function is constructed from the following gray level Stretch fomula:

$$g^{'} = 255 * \frac{g - min}{max - min} \tag{11}$$

The function of gray level Stretching in accordance with each cluster is stated as follows:

$$T (g, lower_V, upwer_V) = 255 * \frac{g - lower_V}{upwer_V - lower_V} \tag{12}$$

where:

min : minimum value.

max : maximum value.

$lower_V$: lower bounds of cluster.

$upwer_V$: upper bounds of cluster.

g : original gray value.

$g^{'}$: new gray value.

With the above way, we constructed the function of gray level adjusting from gray level stretching operator. By the same way, we can construct the functions of gray level adjusting from Hyperbol, Intensification operators... In this study, we construct the functions of gray level adjusting from gray level Stretching and Hyperbol operators in accordance with each cluster and they are listed in Table 2.

Table 2. The functions of gray level adjusting in accordance with each cluster.

Name	Adjusting fomula $T(g)$
Gray level Stretch	$255 * \frac{g - lower_V}{upwer_V - lower_V}$
Hyperbolization	$\left(\frac{255}{e^{-1}-1}\right)\left(e^{-\mu^2(g)} - 1\right)$

Where:

g : original gray value.

$lower_V$: lower bounds of cluster.

$upwer_V$: upwer bounds of cluster.

V : center of cluster.

Values of the thresholds $lower_V$ and $upwer_V$ are calculated automatically in accordance with each cluster V as described in Subsect. 4.1.3.

c-Calculating Automatic Thresholds in Accordance with Each Cluster. Suppose, $d(g)$ is called the distributed function of the g gray level following each cluster. $d(g)$ and the parameters: V, $upwer_V$, $lower_V$ is showed in Fig. 2. The thresholds upper and lower is determined by selecting so that area of subregion between the graphs of $upwer_V$, $lower_V$, the distributed function $d(g)$ and horizontal axis equalize 95

d-Generating Enhanced Image Based on the Gray Level Adjusting in Accordance with Each Cluster. Based on the model of adjusting gray level in accordance with cluster that constructed in Subsect. 4.1.2, each input gray value is changed into new values corresponding to the cluster. This function has general format as follows: $g \mapsto g^{'} = T(g)$. Wherein,

$$T(g) = T(g, lower_V, upwer_V), 0 \leq T(g) \leq 255 \tag{13}$$

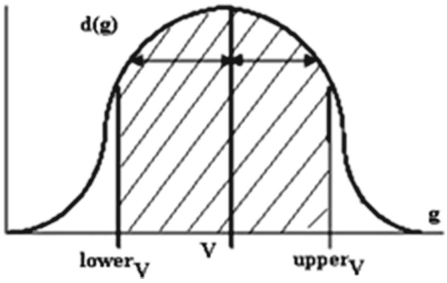


Fig. 2. Distributed function and thresholds following each cluster.

where:
g : original gray value.
$lower_V$: lower bounds of cluster.
$lower_V$: upwer bounds of cluster.
V : center of cluster.

3.2 Upgrading the Algorithm to Multispectral Images

We know that KMeans algorithm clusters the objects which are vectors having many components. So, KMeans algorithm performs well to multispectral images as color images RGB. Therefore, the algorithm LoRSIEK is applied to the multispectral remote sensing images, called LoRSIEK for multispectral images (LoRSIEKMI), as follows:

Step 1: *Localizating input image.* After clustering procedure is executed, each pixel $P\left(g_1^P, g_2^P, g_3^P \ldots\right)$ belong to a cluster. We can consider this is feature of the same pixel relation of the set $\left(g_1^P, g_2^P, g_3^P \ldots\right)$. Thus, overcoming the third disadvantage in Subsect. 3.4.

Step 2: *Constructing gray adjust model.* This step is similar as described in Subsect. 4.1.2.

Step 3: *Calculating automatic thresholds.* The thresholds $lower_V^k$ and $upwer_V^k$ are calculated automatic in accordance with each cluster V and each channel k.

Step 4: *Generating enhanced image.* With each pixel P, we have:

Where: g_k^P : original gray value belong to kth of pixel P. $upper_V^k$: upper bounds of cluster belong to kth channel. $lower_V^k$ lower bounds of cluster belong to kth channel. V : center of cluster. Each center V includes set $(V^1, V^2, V^3 \ldots)$

4 Experiments

We test the proposed algorithm and compare to the fuzzy algorithm.

Data set used for experiments includes 3 types. The first, Landsat ETM+ images are taken in Hoa Binh area in 2001, including 11 pictures about districts and 1 picture about Hoa Binh province. Landsat ETM+ image includes 7 channels [2]: Indigo, Green-red, Red, Near infrared, Medium infrared, Heat infrared, Medium infrared. The second, SPOT images, including 4 channels: Green, Red, Near infrared, Infrared, which are about Hoa Binh and Son La areas with 21 pictures in 2003 and 14 pictures in 2008. In there, SPOT images with high resolution The third, Quickbird images, including 4 channels: Blue, Green, Red and Near infrared, which are downloaded from model data on website: http://opticks.org. Because of the limited scope of the paper, the authors present experiments with different four input images. Results of the proposed method are compared with results of the fuzzy method and result of the recent research of Cheng, in 2003 [19].

To measure the quality of the original and enhanced images, we use the linear index of fuzziness γ [2,8]. Wherein:

$$\gamma = \frac{2}{MN} \sum_{m=1}^{M} \sum_{n=1}^{N} min\left(\mu_{mn}, 1 - \mu_{mn}\right) \tag{14}$$

The index of fuzziness was defined by Kaufmann [14]. The index of fuzziness, for instance, reflects the ambiguity in an image by measuring the distance between its fuzzy property plane and the nearest ordinary plane. This index can be regarded as a degree of difficulty indeciding whether a pixel should be treated as black (dark) or white (bright) [16].

In Experiments, input images clustered into 15 clusters.

4.1 Experiment 1

In experiment 1, original image is a LANSAT image about the Lac Son district, belong to Hoa binh province.

Table 3. Enhancing images with the fuzzy and proposed methods.

Input	Fuzzy Hyp	Fuzzy Int	Cheng	KM Stretch	KM Hyp
					

Table 3 includes input and result images which are enhanced by using the fuzzy method with the intensification and hyperbolization operators and the

proposed method with the stretch and hyperbolization operators. Visually, we see that the enhanced images of fuzzy method do not conserve color correlation, especially with hyperbolization operator. While result images of proposed method still conserve color correlation and prominent than original image. Result of Cheng is better than results of fuzzy method but worse than results of the proposed method.

In Table 4, through comparing fuzzy index of images, we see that γ of result images of the proposed method are smaller than γ of result images of the fuzzy method. This insists that quality of enhanced images of proposed method is higher than quality of enhanced images of fuzzy method.

Table 4. The functions of gray level adjusting in accordance with each cluster.

Input	Fuzzy Hyp	Fuzzy Int	Cheng	KM Stretch	KM Hyp
1	0.34	0.43	0.29	0.18	0.12
2	0.29	0.43	0.32	0.21	0.17
3	0.35	0.40	0.30	0.24	0.20

4.2 Experiment 2

In experiment 2, original image is a LANSAT image about the Lac Thuy district, belong to Hoa binh province.

Table 5 includes input and result images which are enhanced by using the fuzzy method with the intensification and hyperbolization operators and the proposed method with the stretch and hyperbolization operators. Visually, we see that the enhanced images of fuzzy method do not conserve color correlation, especially with hyperbolization operator. While result images of proposed method still conserve color correlation and prominent than original image. Result of Cheng is better than results of fuzzy method but worse than results of the proposed method.

In Table 5, through comparing fuzzy index of images, we see that γ of result images of the proposed method are smaller than γ of result images of the fuzzy method. This insists that quality of enhanced images of proposed method is higher than quality of enhanced images of fuzzy method.

Table 5. Enhancing images with the fuzzy and proposed methods.

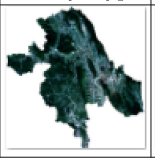
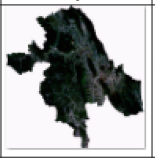
Input	Fuzzy Hyp	Fuzzy Int	Cheng	KM Stretch	KM Hyp
					

Table 6. The functions of gray level adjusting in accordance with each cluster.

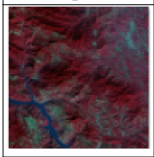
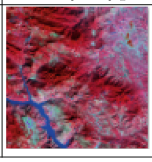
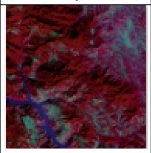
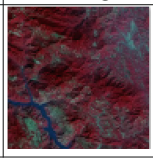
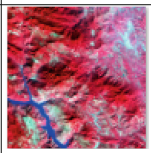
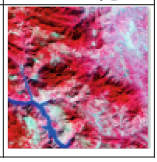
Input	Fuzzy Hyp	Fuzzy Int	Cheng	KM Stretch	KM Hyp
1	0.28	0.28	0.20	0.16	0.11
2	0.27	0.30	0.25	0.21	0.11
3	0.27	0.28	0.23	0.17	0.14

4.3 Experiment 3

In experiment 2, original image is a SPOT image.

Table 6 includes input and result images which are enhanced by using the fuzzy method with the intensification and hyperbolization operators and the proposed method with the stretch and hyperbolization operators. Visually, we see that contrast of the intensification image is only improved slightly from the original, contrast of the hyperbol image is only improved significantly from the original. While result images of proposed method have high contrast. Result of Cheng is better than results of fuzzy method but worse than results of the proposed method.

Table 7. Enhancing images with the fuzzy and proposed methods.

Input	Fuzzy Hyp	Fuzzy Int	Cheng	KM Stretch	KM Hyp
					

In Table 7, through comparing fuzzy index of images, we see that γ of result images of the proposed method are smaller than γ of result images of the fuzzy method. This insists that quality of enhanced images of proposed method is higher than quality of enhanced images of fuzzy method.

Table 8. The functions of gray level adjusting in accordance with each cluster.

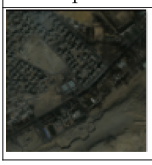
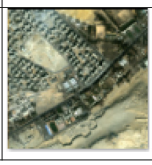
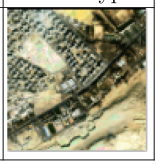
Input	Fuzzy Hyp	Fuzzy Int	Cheng	KM Stretch	KM Hyp
1	0.20	0.50	0.49	0.13	0.08
2	0.43	0.36	0.47	0.32	0.27
3	0.39	0.49	0.60	0.33	0.22

4.4 Experiment 4

In experiment 4, original image is a Quickbird image.

Table 8 includes input and result images which are enhanced by using the fuzzy method with the intensification and hyperbolization operators and the proposed method with the stretch and hyperbolization operators. Visually, we see that the enhanced images of fuzzy method do not conserve color correlation, especially with hyperbolization operator. While result images of proposed method still conserve color correlation and prominent than original image. Result of Cheng is better than results of fuzzy method but worse than results of the proposed method.

Table 9. Enhancing images with the fuzzy and proposed methods.

Input	Fuzzy Hyp	Fuzzy Int	Cheng	KM Stretch	KM Hyp
					

In Table 9, through comparing fuzzy index of images, we see that γ of result images of the proposed method are smaller than γ of result images of the fuzzy method. This insists that quality of enhanced images of proposed method is higher than quality of enhanced images of fuzzy method (Table 10).

Table 10. The functions of gray level adjusting in accordance with each cluster.

Input	Fuzzy Hyp	Fuzzy Int	Cheng	KM Stretch	KM Hyp
1	0.72	0.60	0.47	0.24	0.18
2	0.87	0.84	0.60	0.26	0.22
3	0.85	0.81	0.61	0.32	0.27

5 Conclusions

In this study, we proposed new algorithms of enhancing contrast of remote sensing images LoRSIEK and LoSRSIEMI based on local approach. LoRSIE includes 4 steps. The first, input image is localizated by algorithm KMeans. Then, modifying the gray level adjust formula to fit each cluster. Next, thresholds are calculated automatic in accordance with each cluster. Finally, Generating enhanced image by the gray level model which adjusted in accordance with each cluster. LoSRSIEMI is upgrated for the multispectral remote sensing images. Experiments showed that proposed methods new images which have higher quality than fuzzy method and method of Cheng.

References

1. Bezdek, J.C., Ehrlich, R., Full, W.: FCM: The fuzzy c-Means clustering algorithm. Comput. Geosci. **10**(2-3), 191–203 (1984)
2. Hasanien, A.E., Badr, A.: A comparative study on digital mamography enhancement algorithms based on fuzzy theory. Stud. Inform. Control **12**(1), 21–31 (2003)
3. Zhu, X., Wu, F.: An improved approach to remove cloud and mist from remote sensing images based on Mallat algorithm. In: International Symposium on Photoelectronic Detection and Imaging 2007, Beijing (2007)
4. Bezdek, J.C.: Pattern Recognition with Fuzzy Objective Function Algorithm. Plenum Press, New York (1981)
5. Ross, T.J.: Fuzzy logic with engineering applications. In: Fuzzy Classifying, pp. 379–387. Wiley, Hoboken (2004)
6. Eriksen, J.P., Pizer, S.M., Austin, J.D.: A multiprocessor engine for fastcontrast limited adaptive histogram equalisation. In: SPIE Conference Medical Imaging IV-Image Processing SPIE, vol. 1233 (1994)
7. Gordon, R., Rangayan, R.M.: Feature enhancement of film mammograms using fixed and adaptive neighbourhoods. Appl. Opt. **23**, 560–564 (1984)
8. Sudhavani, G., Srilakshmi, M., Venkateswara Rao, P.: Comparison of fuzzy contrast enhancement techniques. Intl. J. Comput. Appl. **95**(22), 0975–8887 (2014)
9. Zadeh, L.A.: A fuzzy-set-theoretic interpretation of linguistic hedges. J. Cybern. **2**, 4–34 (1972)
10. Pal, S.K., King, R.A.: Image enhancement using smoothing with fuzzy sets. IEEE Trans. Syst. Man Cybern. **SMC-11**(7), 494–501 (1981)
11. Pal, S.K., King, R.A.: On edge detection of X-ray images using fuzzy sets. IEEE Trans. Pattern Anal. Mach. Intell. **PAMI-5**(1), 69–77 (1983)
12. Banks, S.: Signal Processing, Image Processing and Pattern Recognition, Prentice Hall International, Cambridge (1990)
13. Tizhoosh, H.R., Fochem, M.: Image enhancement with fuzzy histogram hyperbolization. In: Proceedings of EUFIT 1995, vol. 3, pp. 1695–1698 (1995)
14. Kaufmann, A.: Introduction to the Theory of Fuzzy Subsets-Fundamental Theoretical Elements, vol. 1. Academic Press, New York (1975)
15. De Luca, A., Termini, S.: A definition of no probabilistic entropy in the setting of fuzzy set theory. Inf. Control **20**, 301–312 (1972)
16. Pal, S.K., Kundu, M.K.: Automatic selection of object enhancement operator with quantitative justification based on fuzzy set theoretic measures. Pattern Recogn. Lett. **11**, 811–829 (1990)
17. Canada Center for Remote Sensing, Fundamentals of Remote Sensing (2008). http://www.ccrs.nrcan.gc.ca
18. MacQueen, J.: Some methods for classification and analysis of multivariate observations. In: Proceedings of the 5th Berkeley Symposium on Mathematical Statistics and Probability, vol. 1, pp. 281–297. University of California Press (1967)
19. Cheng, H.D., Mei, X., Shi, X.J.: Contrast enhancement based on a novel homogeneity measurement. Pattern Recogn. **36**, 2687–2697 (2003)

A Method for Clustering and Identifying HTTP Automated Software Communication

Manh Cong Tran[1(✉)], Hai Nam Nguyen[2], Minh Hieu Nguyen[2], and Yasuhiro Nakamura[1]

[1] Department of Computer Science, National Defense Academy, 1-10-20 Hashirimizu, Yokosuka, Kanagawa, Japan
manhtc@gmail.com, yas@nda.ac.jp
[2] Academy of Cryptography Techniques, 141 Chien Thang, Thanh Tri, Hanoi, Vietnam
nnhai@bcy.gov.vn, hieuminhmta@ymail.com

Abstract. Application developer has trend to take advantage of web as a communication medium environment to reach users because HTTP protocol is mostly allowed in any network environment nowadays. Unfortunately, cyber criminal is also fully exploit HTTP protocol to launch variety of forbidden actions such as application level attacks or spreading malware. Consequently, normal and malicious HTTP automated software (auto-ware) traffic are transparently merged with each other. Clustering and identifying between HTTP communication are raised as serious challenge in order to early investigate internal threats. In this paper, access graph and key features are suggested, based on which HTTP auto-ware communication behavior are recognized. From there, a novelty method in clustering and identifying HTTP auto-ware is presented. Experiment shows promising results since not just malicious communications are detected but also grayware traffic are clustered into groups and identified as their purposes.

Keywords: Automated software · HTTP traffic clustering · Communication identifying · Malware detection

1 Introduction

Application layer attacks have for years posed an ever-serious threat to network security, since they always come after a technically legitimate connection has been established. In a private network, due to security threats, all direct Transmission Control Protocol/Internet Protocol (TCP/IP) outbound/inbound connections should be banned. But HTTP is an exception since, nowadays, application development is transferred more and more onto the web, and everything users need is found through web services. Therefore, in recent years, web is fully utilized as a medium of communication to approach users by spreading HTTP automated software (auto-ware). In consequently, besides of normal users HTTP traffic with their intention, unknown or unnoticed HTTP communications are

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_8

also merged transparently. Unfortunately, these traffic are not just only generated from normal purposes auto-ware such as Operating System(OS) updater but also requested by adware, spyware and bots. Suspicious automated traffic is being recognized as a serious problems since it is able to become internal threats. These matter make the clustering and identifying HTTP auto-ware traffic become a demanding challenge.

To maintain communication, perform updates, or receive commands, all kinds of HTTP-based auto-ware have common characteristics in that they repetitively generate requests, which transparently with normal traffic, to their servers/domains. However, in the details, there are some sophisticated differences in the communications behavior of auto-ware with their sites. In this paper, based on the analysis and study of auto-ware communication behavior, a new method in clustering and identifying of HTTP auto-ware at network level is proposed.

2 Related Work

There were a considerable number of techniques which aim to protect users against malware, however, it continues to be a challenging problem. Many botnets detection methods are presented in [1–3]. Ashley [1] has suggested a method for detecting potential HTTP C&C activity based on repeated HTTP connections to a website. According to this, an algorithm is proposed by for detecting HTTP polling activity. Ashley try to detect the interval in HTTP bot connections, however, it faces with outlier intervals in HTTP bot connections. To overcome this problem, in our method, an outlier intervals recognized algorithm is proposed. Wei Lu et al. in [2], using signature-based techniques, propose a hierarchical framework to automatically discover malicious bot on a large-scale Wi-Fi ISP network, in which the network traffic is classified into different application communities by using payload-signature. These signatures were used to separate known traffic from unknown traffic in order to decrease the false alarm rates. Basil AsSadhan [3] et al. proposed a detection method in which concentrates in C&C communication analysis and find that it exhibits a periodic behaviour. However, these researches [1–3] focus on botnet communication to C&C server, but actually HTTP threats not just come from malicious bots but also can be from other types of automated software such as HTTP spyware, adware or unauthorized applications.

Some of approaches the use lexical features or keywords extracted from URL and web contents as in [4,5]. However, many other type of malicious web pages which are disguised by domain names or URLs like normal website and can harm users PC systems. In this case, lexical or keywords features might be compromised. Bartlett el at [6] proposed an approach to identify low-rate periodic network traffic and changes in regular communication of auto-ware. Their research also focus on many types of auto-ware and monitor TCP flows to detect, but, in our paper, the target not just focus only to detect general types of auto-ware but also on particular URLs where auto-ware request to. In addition, the proposed method just collect and process related HTTP traffic at application layer, this will help reduce expensive in process compare with method use TCP packets process.

3 Features Extraction and Methodology

In target of HTTP traffic observation and analysis from a private network to the Internet, an observation system is implemented. In which, the system will manipulate with all daily outbound HTTP traffic are captured through a proxy server and stored in raw log files.

3.1 Features Extraction

HTTP traffic from a client are constituted many requests from that client to outside. At application layer, a request includes these basic information: IP address of client, full URL, Request method. A full URL's parts contains webpage/server URL and parameter path. At network level, numerous features are extracted which made from basic client requests information, as below.

- Client IP: Source IP address of machine in network which generated requests.
- Request Method: main methods of HTTP requests, POST/GET.
- Request date time: Date and time when a client send request.
- Webpage/Server URL (shorten as URL): URL requested by a Client IP but without parameters part.
- Unique URL: Set of unique URLs requested by a Client.
- Request Interval: Break time between two consecutive requests to the same URLs.
- Request Count: Number of requests to a URL from a client in a period of observation data.
- URL Access Time: a period of time in seconds which a client accessed to a URL from the first request to the end request.
- Client Access Time: a period of time in seconds which a client accessed to the Internet by HTTP protocol. It is determined by checking the difference in seconds between the first and the end request.

3.2 Access Graph

Access graph presents communication behaviour of a client to a specific URL in a duration of time. It is formed on request interval which is extracted from HTTP traffic. Assuming that $R = \{r_1, r_2, ..., r_N\}$ is a set of requests from a client to a server/webpage and all r_i have the same webpage/server URL. Each request r_i is generated at a timing t_i and d_i is interval between request r_i and r_{i+1}, then access graph $G = \{g_1, g_2, ..., g_{N-1}\}$ which including $N - 1$ items $g_i = < t_i, d_i >$. An example of access graph is shown as in Fig. 1, in which, X axis is timing of request and Y axis shows the request interval value in second. An installed auto-ware in a client will behave with different access graph to each URL which is requested to.

3.3 HTTP Auto-ware Communication Behavior

By observation of access graph of many HTTP auto-ware traffic, it can be seen in details that they are subtly different.

Malicious HTTP-based bots always follow the PULL style where they connect to their command and control server periodically in order to get the commands and updates. The number of requests from malicious bots are not high as normal auto-ware (e.g. updater and downloader) which just generate requests with the long interval than unusual malicious bots. An access graph of a malicious auto-ware can be seen in Fig. 1. In that, the main interval is stable at around 900 s. HTTP gray auto-ware which used for updating news or advertising contents will access to multiple URLs. As a result, if two URLs which requested by the same a auto-ware, they will have the similar behaviour and access graph. These features are different from malicious bots when they often connect to one control domain and to a specific server resource. In addition, many URLs are requested with the same timing by a specified auto-ware, so the access duration to these URLs is absolutely equal. It means that the first and the last requests timing to these URLs are the same with others. In Fig. 2 shows an example of access graph of an auto-ware from a client. In which, event this auto-ware accesses to difference two URLs but similar access graphs are behaved, and the first and the last requests moment of them are equal. Suspicious auto-ware, such as adware, trends to update advertising contents from many different sites which are not just from the same top level domains. This matter is different with news or normal update auto-ware since they mostly communicate to the same top level domain.

Auto-ware, especially malicious bots are always activated, communicate to their sites or servers, as soon as computer or system is turned on. That means the access time of auto-ware is almost the same with working time of computer or system.

On contrary with auto-ware, there are no interval or periodic pattern in users' web access. However, in recent years, many sites (e.g. online shopping, social

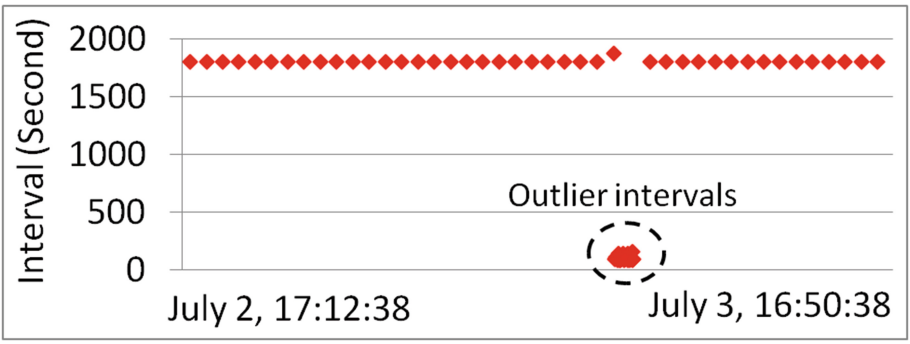


Fig. 1. An access graph of an malicious bot to a URL

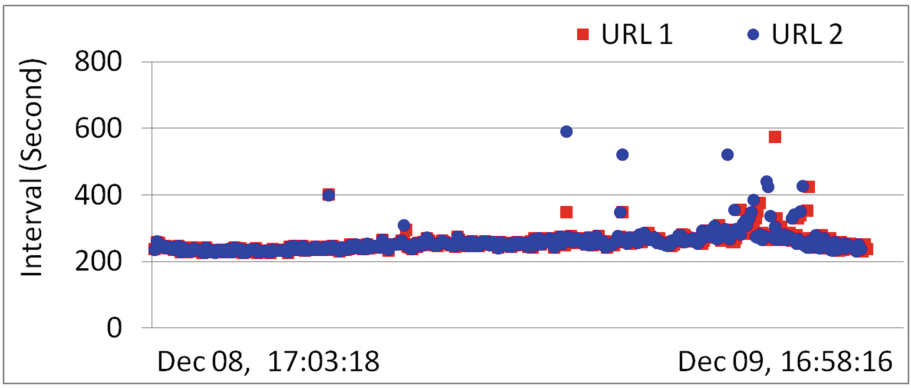


Fig. 2. Similar access graphs from an auto-ware of a client IP to two different URLs, and the first and the last requests moment of these URLs are equal, from Dec 08 17:03:18 to Dec 09, 16:58:16

media webpage) append advertisement path to their sites and use JavaScript or Flash as auto-aware part to automatically collect the advertising content as adware or spyware. Therefore, parts of users access sites can act as auto-ware however the automated accessing time from users traffic is not much comparing to working time of computer or system.

4 Clustering and Identifying Proposed Method

Based on the auto-ware communication behavior which is analyzed via the observation of access graphs in Sect. 3, a clustering and identifying of HTTP auto-ware communication method, including four phrases, is proposed as in Fig. 3, details as follow.

4.1 Pre-processing Phase

This preprocessing phase is objective to eliminate unnecessary processed data. For each client IP, the one day HTTP traffic features are extracted and preprocessed, in order to process in this phase two methods are applied:

The first one is to filter URLs requests from Client IP through a white list of second level domain name (SLDN). This filter method is described in [4], according to that, the tokens in the URLs of phishing websites are less consistent with their content when compared with those of legal websites. In normal way, the legitimate website contains the brand or company names in the SLDN. Even though the phishing website also contains these keywords in the URL but it can not be in the SLDN. Therefore, a domain name which contains a second level domain name is defined in SLDN white list is marked as benign. The second method is based on the number of requests to a URL from a client IP.

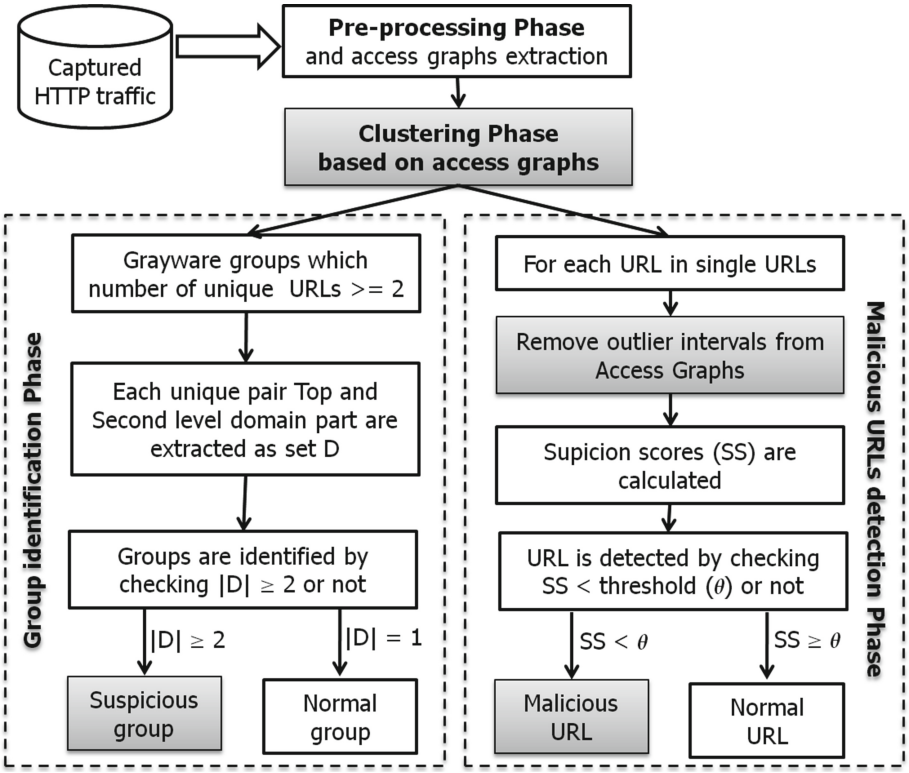


Fig. 3. Clustering and identifying proposed method flowchart

Suspicious auto-ware have trend to access many times to a URL in a duration of time. Therefore, if the number of requests to a URL is too small, it seem not to request by an auto-ware.

4.2 Clustering Phase

After pre-processing phase, in this phase, remaining URLs will be clustered into number of groups based on their characteristics which are presented in Sect. 3. Accordingly, two URLs are the same group (requested by the same auto-ware from a client) if they match one of the following conditions:

- The first and the last request timing to two URLs are the same.
- Based on the similarity of its access graph, if they are small enough, they will be marked as in the same group.

In order to score how similarity between any two access behaviour, Modified Hausdorff (MH) distance, which is presented in [7], is suggested. Assuming that there are two access graphs $A = (a_1, ..., a_N)$ and $B = (b_1, ..., b_M)$.

Defining that the distance between two points a_i and b_j is calculated as Euclidean distance $d(a_i, b_j) = \|a_i - b_j\|$. From that, distance between a point a_i and graph B is defined as $d(a_i, B) = min_{b_j \in B} \|a_i - b_j\|$. Generalized Hausdorff distance of A and B in [7] is defined as below:

$$d(A, B) = \frac{1}{N} \sum_{a_i \in A} d(a_i, B) \tag{1}$$

Based on (1), distance between access graph A and B, which follow by MH distance (MHD) is formed as below:

$$MHD(A, B) = max(d(A, B), d(B, A)) \tag{2}$$

The smaller MH distance between A and B, the more A and B are similar each other. After this phase, grayware is clustered into groups which contain at least two unique URLs. The others single URLs are not belong any group. From there, two follow phases in order to identify grayware group and detect malicious URL are processed.

4.3 Greyware Identification Phase

After clustering phase, groups which contain 2 URLs above will be identified into two type of group normal or suspicious. The steps for this phase is shown in the left branch of Fig. 3. This phase is based on the analysis in Sect. 3.3, in which, suspicious auto-ware trends access to many different domains to maintain advertising contents. Therefore, in this phase, by count the unique domains in each group, if it greater or equal 2 group will be marked as suspicious group. Vice versa, group will be marked as normal group.

4.4 Malicious URLs Detection Phase

For URLs can not be clustered, a suspicion score based on their access graphs is presented to recognize a URL is malicious or normal. As analysis in Sect. 3.3, malicious bot always communicate to a specific URL or resource by automatically generating requests with a stable interval, therefore, malicious bot access graph almost has no variation. In this phase, the method do not try to detect the interval of malicious requests but in target to score the variation of access graph. However, the interval of malicious bot requests are changed some times, these are outlier intervals, but the main interval is steady. As can be seen in Fig. 1, some intervals are uncertain but stable interval is around 1800 s. In order to eliminate outlier intervals from access graph X, a nested loop algorithm is proposed and described in Table 1.

However, The removed intervals must not change the main shape of access graph, in other words, access behavior is not effected. This matter is determined by checking the difference of total intervals before and after each time removing outlier values. By implementation in this work, the value 0.6 is used as threshold

Table 1. Nested loop to eliminate outlier intervals from a access graph X

Step	Description
1	Calculate total access time from client to R, $S = x_1 + x_2 + .. + x_N$
2	Determine the average $\bar{x}$ of X
3	Determine maximum distance d from x_i to $\bar{x}$, denoted $dmax = max(d_1, .., d_i, .., d_N)$, which $d_i = \|x_i - \bar{x}\|$
4	Remove all x_j, which has $d_j = dmax$, from X. After this step, number of items in X will be decrease and become X', denote that $M = \|X'\|$
5	Calculate $S' = x_1 + x_2 + .. + x_M$, x_i is an element in X'. if $S' < S * 0.6$ then go to end, else repeat the loop from Step 2

in Step 5 is suggested. Overall of this phase can be seen in the right branch of Fig. 3, in which before suspicion score is determined the outlier intervals removal is processed. Suspicion score of a URL is calculated as *coefficient of variation* of access graph, which define as is based on deviation and mean of access graph. The smaller suspicion score shows that URL is more suspicion.

Table 2. Results of URLs clustering and grayware groups identification

Log no	Pre-processing phase		Clustering phase		Clustered groups identification					
	Captured request number	Unique URLs remaning	Single URLs	Total groups	Normal groups			Suspicious groups		
					Sub total	True (NT)	False (NF)	Sub total	True (ST)	False (SF)
1	197,671	173	39	33	10	6	4	23	22	1
2	89,336	5	0	1	1	0	1	0	0	0
3	354,113	14	5	2	1	1	0	1	1	0
4	113,885	10	2	3	3	2	1	0	0	0
5	98,344	35	0	10	6	5	1	4	3	1
6	83,063	262	26	42	6	3	3	36	33	3
7	65,263	18	5	4	4	2	2	0	0	0
8	44,910	68	1	6	6	6	0	0	0	0
9	40,190	289	15	39	26	17	9	13	12	1
10	18,359	21	7	4	2	0	2	2	1	1
11	28,633	126	7	18	15	14	1	3	1	2
12	66,287	46	17	4	0	0	0	4	3	1
13	16,925	11	7	1	0	0	0	1	1	0
14	38,429	139	4	23	12	10	2	11	8	3
15	40,754	226	13	27	12	9	3	15	12	3
16	22,305	15	8	3	2	0	2	1	1	0
Total	1,318,467	1,458	156	220	106	75	31	114	98	16
Propotion (%)					**48.18**	**34.09**	**14.09**	**51.82**	**44.55**	**7.27**
Accuracy ((NT + MT)/Total group)							**78.64 %**			
Error rate ((ST + SF)/Total group)							**21.36 %**			

5 Experimental Results and Discussion

For experiment purpose, 16 log data of HTTP outbound traffic which are captured from a private network. One log data contain HTTP traffic in one day of a client IP. They are used as input of system which is implemented as flowchart in Fig. 3. In order to evaluate the method, after results are archived from experimental system, it will be checked by two system in [8,9]. In that, [8] is a online service and [9] is installed in the experimental network.

After pre-processing and clustering phase, clustered results are summarized in Table 2. After pre-processing phase by eliminating normal URLs using the method as described in Sect. 4.1, it can be seen that almost clients are accessed to normal URLs since not many unique URLs are remained. After clustering phase, almost URLs are clustered into grayware groups with a proportion of 89.3 % since only 10.7 % number of URLs (156 URLs) are not clustered in any groups. There are six clients has remained greater or equal 126 URLs and these clients look more suspicion since the number of grayware is higher than others.

After first two phases, grayware identification phase which is described in Sect. 4. By manually checked with the support of [8,9], the detail of identification accuracy results are also summarized in Table 2. Accordingly, the identification accuracy constitutes a reasonable rate 78.64 %, in which false negative rate (14.09 %) is higher than true negative rate (7.27 %). Remaining single URLs

Table 3. Malicious URLs detection results

Log no	Total URLs	Normal			Malicious		
		Sub total	True (NT)	False (NF)	Sub total	True (MT)	False (MF)
1	39	39	35	4	0	0	0
2	0	0	0	0	0	0	0
3	5	5	5	0	0	0	0
4	2	2	1	1	0	0	0
5	0	0	0	0	0	0	0
6	26	26	24	2	0	0	0
7	5	5	5	0	0	0	0
8	1	1	1	0	0	0	0
9	15	11	11	0	4	3	1
10	7	4	4	0	3	3	0
11	7	4	4	0	3	3	0
12	17	13	11	2	4	4	0
13	7	5	4	1	2	2	0
14	4	1	1	0	3	3	0
15	13	11	11	0	2	2	0
16	8	6	6	0	2	2	0
Total	**156**	**133**	**123**	**10**	**23**	**22**	**1**
Propotion		**85.26 %**	**79.85 %**	**6.41 %**	**14.74 %**	**14.10 %**	**0.64 %**
Accuracy ((NT + MT)/Total Urls)					**92.95 %**		
Error rate ((NF + MF)/Total Urls)					**7.05 %**		

become input data of malicious URLs detection phase. The results and detection accuracy are included in Table 3. In this table, a promised detection rate is shown since accuracy rate constitutes 92.95 %. However, error rate still contributes 7.05 %. In this phase, false negative rate (6.41 %) is also higher than true negative rate (0.64 %). By manually investigation, error rate is caused by some URLs of grayware is communicated by separated behavior therefore they are not clustered into any group in clustering phase.

6 Conclusion and Future Work

In this paper, a new clustering and identifying HTTP auto-ware communication method is presented. In which, various types of HTTP auto-ware communication behavior is analyzed based on access graph. From there, HTTP traffic are clustered, detected and identified by their purposes. The results are promising although error rate still is contributed by some reasons which discussed in Sect. 5. The future work will concentrate to improve the clustering algorithm which is considered to reduce the false negative rate. Moreover, due to a lot of HTTP traffic with secure connections, the method need to be improved to clarify not just HTTP communication but also HTTPs traffic.

References

1. Ashley, D.: An algorithm for http bot detection. University of Texas at Austin - Information Security Office (2011)
2. Lu, W., Tavallaee, M., Ghorbani, A.A.: Automatic discovery of botnet communities on large-scale communication networks. In: Proceedings of the 4th International Symposium on Information, Computer, and Communications Security, pp. 1–10. ACM, Sydney (2009)
3. AsSadhan, B., Moura, J.M.F.: An efficient method to detect periodic behavior in botnet traffic by analyzing control plane traffic. J. Adv. Res. **5**(4), 435–448 (2014)
4. Chen, Y.-S., Yu, Y.-H., Liu, H.-S., Wang, P.-C.: Detect phishing by checking content consistency. In: 2014 IEEE 15th International Conference on Information Reuse and Integration (IRI), pp. 109–119. IEEE, Redwood City, August 2014
5. Chen, T.-C., Dick, S., Miller, J.: Detecting visually similar Web pages: application to phishing detection. ACM Trans. Internet Technol. **10**(2), 1–38 (2010). Article 5. ACM
6. Bartlett, G., Heidemann, J., Papadopoulos, C.: Low-rate, flow-level periodicity detection. In: 2011 IEEE Conference on Computer Communications Workshops (INFOCOM WKSHPS), pp. 804–809. IEEE, Shanghai, April 2011
7. Dubuisson, M.-P., Jain, A.K.: A modified Hausdorff distance for object matching. In: 1994 Proceedings of the 12th IAPR International Conference on Pattern Recognition, vol. 1 - Conference A: Computer Vision & Image Processing, vol. 1, pp. 566–568. IEEE, Jerusalem, 9–13 October 1994
8. Virus Total. http://virustotal.com/. Last checked on May 2016
9. McAfee Web Gateway. http://www.mcafee.com/us/products/web-gateway.aspx. Last checked on May 2016

A New Neuro-Fuzzy Inference System for Insurance Forecasting

Le Hoang Son[1(✉)], Mai Ngoc Khuong[2], and Tran Manh Tuan[3]

[1] VNU University of Science, Vietnam National University, Hanoi, Vietnam
sonlh@vnu.edu.vn
[2] Post and Telecommunication Insurance, Hanoi, Vietnam
khuongmaingoc@gmail.com
[3] University of Information and Communication Technology, Thai Nguyen, Vietnam
tmtuan@ictu.edu.vn

Abstract. Insurance forecasting is a matter of vital importance to insurance companies for analyzing of annual income, premium and loss reserving, loss payment, etc. Recent years have also seen increasing discussion within the actuarial community of the need for insurance forecasting techniques that are more solidly grounded in rigorous machine learning methodologies. Taking advantages of knowledge-reuse and learning capability for dealing with uncertainties, hybridization of neural networks and fuzzy logic could enhance the accuracy of forecasting for insurance applications. In this paper, we propose a novel neuro-fuzzy inference system for insurance forecasting. It uses multiple parameter sets where each set is responsible for a small subset of records. The aim of each parameter set is to minimize Mean Square Error within records of the subset. The learning strategy and a rule reduction method are also proposed. Empirically validation on the benchmark and real insurance datasets show the advantages of the new system.

Keywords: Accuracy · Insurance forecasting · Learning strategy · Neuro-fuzzy systems · Rule reduction

1 Introduction

Insurance forecasting is one of the major challenges for insurance companies including the prediction of annual income, premium and loss reserving, loss payment, etc. [2] Specifically, predicting future losses is done through an analysis of past loss data, which usually span a sufficient number of years to achieve some degree of credibility [8]. The law of large numbers, exposure data, any anticipated changes in company operations or structure, inflation, workers compensation benefit changes, and any other relevant factors must be considered when forecasting losses [16,18]. Adjusting historical losses to account for inflationary trends so that their value is in current dollar amounts is the main task of loss trending forecast [9]. Historical loss amounts are multiplied by trending factors to convert historical loss amounts to current dollar amounts [9].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_9

This kind of estimation is therefore a matter of vital importance to the company [21]. In discussing loss reserve that was experienced by the insurance industry, the actuarial profession faces with many failures in prediction of loss reserving [7]. Recent years have seen increasing discussion within actuarial community of the need for insurance forecasting techniques that are more solidly grounded in rigorous machine learning methodologies [5–7].

Claims and risks have long been estimated using a pure algorithmic technique or a simple stochastic technique [5–7,14]. Zhang, Dukic and Guszcza [22] proposed a Bayesian non-linear hierarchical model that addresses some of the major challenges of non-life insurance companies when forecasting the outstanding claim amounts for which they will ultimately be liable. Bernoth and Pick [10] proposed an empirical framework that allows estimating unobserved linkages in panel data sets that contain observed regressors for forecasting the fragility of the banking and insurance sectors. Martínez-Miranda et al. [13] argued that better statistical properties can be obtained when more aggregated data is available for the statistical analysis than just the classical aggregated payments. Ana-Maria and Ghiorghe [4] analyzed and forecasted the evolution of the marine insurance market through autoregressive models. Marine insurance facilitates global trade, ensures economic property, provides peace of mind, improves quality of life and provides social benefits. Kaymak [19] conducted an exploratory research towards the application of collaborative Markov chains as a forecasting model within the field of healthcare insurance. The model was based on both predicting care demand and associated institutional pathway traversal [19]. Taking advantages of knowledge-reuse and learning capability for dealing with uncertainties, Peter Mulquiney et al. [17] used Artificial Neural Networks (ANN) to estimate the future payments for the claims which have occurred on an insurance portfolio. By supplementing their ANN with separate forecasts which accounted for the expected changes in the future claims environment, they have addressed a difficulty in forecasting future payments is that the time series of payments often depends on influences that are not observable in the historical data [17]. Shapiro [1,2] reviewed fuzzy logic applications in insurance so as to document the unique characteristics of insurance as an application area. Abdullah and Rahman [11] presented a decision model using Fuzzy Inference System (FIS) to identify the likelihoods of purchasing health insurance based on the selected risk factors. Three risk factors were considered as the input of the system including age, salary and risk of having illness. The likelihoods of purchasing health insurance was the output of the system and defined in three linguistic terms of 'Low', 'Medium' and 'High' [11]. Input and output data were governed by the Mamdani inference rules of the system to decide the best linguistic term. The usage of fuzzy inference system would offer possible justifications to set a new approach in identifying prospective health insurance purchasers [11]. Utilizing neural networks and fuzzy logic for insurance applications, e.g. ANFIS and MANFIS, was shown in [3,12,15].

It has been observed that the hybridization of neural networks and fuzzy logic could enhance the accuracy of forecasting for insurance applications [3].

Therefore, in this paper, we propose a novel neuro-fuzzy inference system for insurance forecasting. It uses multiple parameter sets where each set is responsible for a small subset of records. The aim of each parameter set is to minimize Mean Square Error (MSE) within records of the subset. The learning strategy and a rule reduction method are also proposed. The new system is empirically validated on the benchmark UCI Machine Learning Repository and real insurance datasets. The rest of the paper is organized as follows. Section 2 describes the proposed approach. Section 3 validates the method on the benchmark UCI and real insurance data. Finally, conclusions and further works are covered in last section.

2 A New Neuro-Fuzzy Model for Insurance Forecasting

In this section, we present the architecture of the system, learning methods in both training and testing phases, and the rule reduction in sub-sections accordingly.

2.1 Design

The proposed system has 2 main parts: fuzzy system (FS) and neural network (NN) (Fig. 1). FS consists of 3 layers:

– Layer 1 (fuzzification layer): calculates a membership degree for each crisp value from a given record α_i $(i = 1, ..., N)$ where the functions can be the trapezoid, triangular, Gaussian, etc. depending on a specific problem.

$$\mu_{Aj}(x_1), \mu_{Bj}(x_2) \tag{1}$$

– Layer 2 (rule layer): combines membership degrees to generate strength of individual rule.

$$w_i = \Gamma(\mu_{Aj}(x_1), \mu_{Bj}(x_2)) \tag{2}$$

where $\Gamma(.)$ represents an inference function, e.g. product or min operators.

– Layer 3 (normalization layer): normalizes strength of each rule provided that summation of all rule strengths of one data equals to 1.

$$\overline{w}_i = \frac{w_i}{\sum\limits_i w_i} \tag{3}$$

A NN consists of 3 layers:

– Layer 1 (transfer layer): transfers input data to the next layer.

– Layer 2 (consequent layer): computes consequent part of each rule from the input.

$$f_i = a_i x_1 + b_i x_2 + cz \tag{4}$$

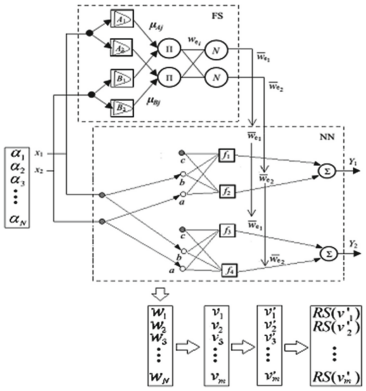


Fig. 1. The architecture of a neural network

– Layer 3 (output layer): mix consequent values with the rule strengths to generate final outputs.

$$y_1 = \overline{w}_1 f_1 + \overline{w}_2 f_2 \; and \; y_1 = \overline{w}_1 f_3 + \overline{w}_2 f_4 \tag{5}$$

The architecture of a neural network is presented as in Fig. 1 below.

It is definitely that a set of parameters including those in equations (1, 4) is trained by a leaning algorithm so that the MSE value for this record is minimal. We denote this set as v. However, when dealing with other records α_i $(i = 1, ..., N)$, using a set v only would make it tend to lean to the last record.

In order to handle this issue, the proposed system uses the representative set $V = v_1,\ v_2,\ ...,\ v_k$ where each $v_i,\ (i = 1, ..., k)$ is responsible to calculate the MSE values for some related records in $(\alpha_1, \alpha_2, ..., \alpha_N)$. Using the representative set of parameters would both make the accuracy better than using a parameter set only and ensure the computational complexity. This is the trade-off between the accuracy and the computational time.

2.2 Training

In order to train the representative set, we propose the learning strategy including global and local training phases. In the local training, we use gradient descent to get the optimal parameter set.

In global training, we start from a random parameter set w_0 and train it by local training to get the optimal w_1 for the first record. Next, compare w_0 and w_1. If they are far enough then a set of representatives consisting of w_0 and w_1 is formed. This set is then applied to the second record by plugging each parameter set for the record and deciding the most suitable parameter set based on their MSE value. Notice that this process is used to determine the initial parameter set for second record from the set of representatives only, and no further local

training is required. Then, we continue to use the local training for second record to get the optimal w_3. Again, the comparison between w_3 and the current set of representatives is performed, and if w_3 is far enough to all representatives then it is put into the set. The process is repeatedly performed until the last record and we will get the final optimal set of representatives re-denoted as $V = v_1,\ v_2,\ ...,\ v_k$. The global training has 4 main steps:

Step 1: Normalize the dataset.

- Normalize all inputs to avoid bias to parameters that are corresponding to the inputs.
- Normalize all outputs to avoid bias to any output.

$$y*_{1i} = \frac{y^0_{1i} - min_i(y^0_{1i})}{max_i(y^0_{1i}) - min_i(y^0_{1i})} \tag{6}$$

$$y*_{2i} = \frac{y^0_{2i} - min_i(y^0_{2i})}{max_i(y^0_{2i}) - min_i(y^0_{2i})} \tag{7}$$

Step 2: Train the whole system to generate W.

- From the first record, use local training to generate parameter.
- Repeat step 2.1 for other data to generate the full set of parameter vectors $W = \{w_i, i = \overline{1, N}\}$.

Step 3: Reduce the full set to generate the representative set V.

- Assign w_1 to v_1.
- Assume we have $v_1,\ v_2,\ ...$, and the current record generates w_{cur}. If there exists $v_j : |v_j - w_{cur}| < d$, then we drop w_{cur}. It not, then w_{cur} will be put into the representative set. Repeat this step until we scan all records and get the representative set: $V = v_1,\ v_2,\ ...,\ v_k$.

Step 4: Reduce the number of rule sets for each group.
Step 5: Training the representative set $V = v_1,\ v_2,\ ...,\ v_k$ one more time to make the final representative set $V' = v'_1,\ v'_2,\ ...,\ v'_k$ that best fits the model.

2.3 Testing

After the training phase, we have the representative set. Thus, with a new testing record, we would have no idea what group it belongs to. To know the group of a new data record, we simply solve the classification problem. In this case, we use FKNN (Fuzzy K-Nearest Neighbors) algorithm due to its simplicity and straightforward.

FKNN is a time-consuming algorithm because it must scan the whole training data to find K-nearest neighbors. This is the down point of the method compared to other classic models, because other models generate only one parameter set and they do not need to do anything but apply directly that set to testing record. In the experimental part, we will see how much time this phase consumes compared to other models.

2.4 Rule Reduction

Being noted from the learning scheme that it is necessary to create a subset of the original fuzzy rules for each group of records. Each record after the local training phase, will have a set of weights and each weight is regarded as the firing strength of a rule to that record. Denote the set of weights as $w^{(ir)},\ r = \overline{1,R}$, i is index of a record, r is index of a rule in the rule sets. R is the number of rules in the complete rule sets. After dividing all records into different groups, assume that each group has a suitable rule sets, then associated rules of a specific group will have larger values of weights than those that do not belong to that group. Consider a group A, which N_A has data records, its rule weight is $w_A^{(ir)}, r = \overline{1,R}$. Rearrange the weights with respect to index r, we gain $w_A'^{(ir)}$ satisfying $w'^{(ir)} \geq w_A'^{(ir')}$ provided that $r \geq r'$. Pick up candidate rules for individual record from larger weights to smaller ones with the constraint summation of weights of picked rules contribute a large amount in summation of weights of all rules.

$$\frac{\sum\limits_{r=1}^{p} w_A'^{(ir)}}{\sum\limits_{r=1}^{R} w_A'^{(ir)}} \geq T, \tag{8}$$

where p is the position we pick up the first rule that satisfies the above condition. T is some threshold which is close to 1. For example 0.9 or 0.95. Now, the task is to pick a rule set which is close to all other candidate rule sets of all data records. Mark the unarranged rule weights $w_A^{(ir)}$ as 1 indicating candidate rules and 0 otherwise. Then each record will have a binary vector representing the candidate rule sets. Intuitively, we can use the Jaccard index to measure the difference between two rule set. Define Jaccard dissimilarity measure between 2 sets A and B as follow:

$$d(A,B) = 1 - \frac{|A \cap B|}{|A \cup B|} \tag{9}$$

It is easy to recognize that Jaccard dissimilarity is a metric distance. We can pick a rule set for group A that fits the following condition:

$$RS(A) = \{RS_i | \sum_{k \neq i} d^2(RS_i, RS_k) = min(\sum_{k \neq i} d^2(RS_i, RS_k))\} \tag{10}$$

$RS(A)$ is the reduced rule set of group A, RS_i is candidate rule set of the i^{th} record. By this method, we only care about the similarity with respect to candidate rules. The following method is much faster and depends on non-candidate rules. Denote the binary vector of candidate rule set as $B_i = \{b_{ij}\}, i = \overline{1,R}$ where b_{ij} is 0 indicating non-candidate and 1 indicating candidate-rule. The distance or dissimilarity measure is simply computed as follow:

$$d(RS_i, RS_j) = \sqrt{\sum_{k=\overline{1,R}} (b_{ik} - b_{jk})^2} \tag{11}$$

Denote the rule set we are looking for as $X = \{x_j\},\ j = \overline{1, R}$ the constraint to pick the best X is nearly the same to formula (10).

$$RS(A) = \{X| \sum_{i=\overline{1,N_A}} d^2(X, RS_i) = min(\sum_i d^2(Y, RS_i))\} \tag{12}$$

The value of X is not limited in the available rule sets of group A, but it can get any possible value. Solve the problem in (12) and get the solution:

$$x_j = \begin{cases} 0,\ if \sum\limits_i b_{ij} < N_A \\ 1,\ if \sum\limits_i b_{ij} \geq N_A \end{cases} \tag{13}$$

3 Experiments

In this section, we implemented the proposed neuro-fuzzy system denoted as MANFIS-S in addition to the relevant neuro-fuzzy models for insurance forecasting namely ANFIS [15] and MANFIS [12] in Matlab. In order to evaluate

Table 1. MSE of training and testing phase using K-fold (K = 3)

Model	Fold	Data 1		Data 2	
		Training	Testing	Training	Testing
ANFIS	1	7.95E-03	0.0032	0.0601	0.0448
	2	6.50E-03	0.0126	0.0529	0.0591
	3	7.55E-03	0.0059	0.0592	0.0465
	Avg	7.30E-03	0.0072	0.0574	0.0501
	Sig	3.70E-07	1.56E-05	1.03E-05	4.07E-05
MANFIS	1	0.0222	0.0021	0.0603	0.0446
	2	0.0448	0.0516	0.053	0.0594
	3	0.0189	0.0138	0.0594	0.0463
	Avg	0.0286	0.0225	0.0576	0.0501
	Sig	1.33E-04	4.46E-04	1.06E-05	4.37E-05
MANFIS-S1	1	5.60E-03	0.0025	8.87E-05	0.0481
	2	2.30E-03	0.0106	6.39E-05	0.0638
	3	3.90E-03	0.0064	6.61E-05	0.0484
	Avg	3.90E-03	0.0065	7.29E-05	0.0534
	Sig	1.82E-06	1.09E-05	1.25E-10	5.37E-05
MANFIS-S2	1	8.20E-03	0.0032	2.77E-05	0.0478
	2	7.70E-03	0.0114	2.80E-05	0.0637
	3	1.56E-02	0.0148	2.33E-05	0.0485
	Avg	1.05E-02	0.0098	2.63E-05	0.0533
	Sig	1.30E-05	2.37E-05	4.69E-12	5.38E-05

the accuracy, Mean square error (MSE) is used.

$$MSE = \frac{\sum_{i=1}^{n}(M_{corrected}(i) - M_{predicted}(i))^2}{n} \quad (14)$$

where $M_{corrected}(i)$ and $M_{predicted}(i)$ denotes the correct and predicted values, respectively. The experiments are run on the system with configuration of 2 G RAM, 2.13 GHz core 2 Duo. The experimental data include a real insurance dataset from Post and Telecommunication Insurance (PTI), Vietnam (Data 1) and a benchmark UCI Machine Learning dataset [20] (Data 2).

- Data 1: contains 10 input components and compensation via more than 100 insurance operations and over 26 data records corresponding to 26 companies in PTI. 182 insurance operations are PTI's services which are spread on 4 main product groups: motor vehicle insurance, personal insurance, property insurance and marine insurance. Two outputs are the predictions of revenue and average compensation of PTI insurance in the next year;
- Data 2: contains 17 input components over 5000 data records. This is a small version of the original dataset for the sake of computational complexity.

We also implemented two variants of MANFIS-S namely MANFIS-S1 (MANFIS-S with the rule reduction strategy) and MANFIS-S2 (MANFIS-S

Table 2. Comparison of computational time in training and testing phase

Model	Fold	Average time for 1 iteration (sec)			
		Data 1		Data 2	
		Training	Testing	Training	Testing
ANFIS	1	0.01723	9.85E-04	29.7	5.16E-04
	2	0.01718	2.22E-03	28.4	4.90E-04
	3	0.01714	5.96E-04	28.7	5.23E-04
	Avg	0.0172	1.27E-03	28.9	3.53E-04
MANFIS	1	9.40E-03	3.18E-04	28.25	5.14E-04
	2	9.34E-03	3.03E-05	28.15	4.98E-04
	3	9.76E-03	4.70E-04	29.1	5.13E-04
	Avg	9.50E-03	2.73E-04	28.5	5.08E-04
MANFIS-S1	1	0.03933	1.62E-03	46.83	0.0184
	2	0.03096	9.32E-04	45.41	0.0182
	3	0.02165	9.18E-04	43.67	0.0177
	Avg	0.0306	1.16E-03	45.3	0.0181
MANFIS-S2	1	0.02874	1.03E-03	59	0.0192
	2	0.03872	1.25E-03	56.8	0.019
	3	0.02994	1.18E-03	57	0.0188
	Avg	0.0325	1.16E-03	57.6	0.019

without the rule reduction). The aim is to validate the efficiency of the rule reduction method. In the following tables, Avg is abbreviation for the average value and sig is standard deviation.

Table 1 show the average MSE and standard deviation values of all algorithms on the training and testing phases, respectively. It indicated that the proposed models (MANFIS-S1 and MANFIS-S2) have better accuracy than MANFIS and ANFIS. It is obvious that MSE values of MANFIS-S1 are the smallest among all.

Table 2 shows the computational time of each model through the average time for each iteration in the training and testing phases on both datasets. It has been validated that both MANFIS-S models consume approximately two times longer than ANFIS and MANFIS take.

4 Conclusion

This paper proposed a new neuro-fuzzy model for insurance prediction. The main points from prediction of insurance exercise were the use of a representative set of parameters which could make high accuracy and ensure the computational complexity. A new training strategy and especially a new method to reduce the number of rules for each group were proposed and embed in the model. The training combines global and local phases in which the global training acts through the representative problem and the local one uses gradient descent for utilizing the advantages of simplicity, the computational time, and accuracy of the system. The new system has been empirically validated, and the experimental results demonstrated its advantages over the relevant ones.

A difficulty in insurance forecasting is quantities of information gathering. Compensation information contains only one value of loss payment and lots of information is not observable because of input data negligent exploitation from staff. Thus, further works of this research aim to handle this problem by the following directions. Firstly, an idea of dynamic neuro-fuzzy models in terms of dynamic inputs and outputs is investigated. Secondly, convergence rate analysis in some special cases of datasets should be performed for the complete evaluation of the model. Lastly, applications to other specific prediction tasks for insurance should be investigated.

Acknowledgement. The authors would like to thank the Center for High Performance Computing, VNU University of Science for executing the program on IBM Cluster 1350 server.

References

1. Shapiro, A.F.: An overview of insurance uses of fuzzy logic. In: Chen, S.-H., Wang, P.P., Kuo, T.W. (eds.) Computational Intelligence in Economics and Finance, pp. 25–61. Springer, Heidelberg (2007)
2. Shapiro, F.: Fuzzy logic in insurance. Insur. Math. Econ. **35**(2), 399–424 (2004)

3. Shapiro, F.: Insurance applications of neural networks, fuzzy logic, and genetic algorithms. In: Intelligent and Other Computational Techniques in Insurance: Theory and Applications (2003)
4. Ana-Maria, B., Ghiorghe, B.: Application of autoregressive models for forecasting marine insurance market. Ovidius Univ. Ann. Ser. Econ. Sci. **13**(1), 1125–1129 (2013)
5. De Alba, E., Nieto-Barajas, L.E.: Claims reserving: a correlated Bayesian model. Ins. Math. Econ. **43**, 368–376 (2008)
6. De Alba, E.: Bayesian estimation of outstanding claims reserves. N. Am. Act. J. **6**(4), 1–20 (2002)
7. De Alba, E.: Claims reserving when there are negative values in the runoff triangle: Bayesian analysis using the three-parameter log-normal distribution. N. Am. Act. J. **10**(3), 1–15 (2006)
8. Gaver, J.J., Paterson, J.S.: Do insurers manipulate loss reserves to mask insolvency problems? J. Acc. Econ. **37**, 393–416 (2004)
9. Antonio, K., Beirlant, J.: Issues in claims reserving and credibility: a semiparametric approach with mixed models. J. Risk Ins. **75**, 643–676 (2008)
10. Bernoth, K., Pick, A.: Forecasting the fragility of the banking and insurance sectors. J. Bank. Finance **35**(4), 807–818 (2011)
11. Abdullah, L., Rahman, M.N.A.: Employee likelihood of purchasing health insurance using fuzzy inference system. Int. J. Comput. Sci. **9**(1), 112–116 (2012)
12. Son, L.H., Linh, N.D., Long, H.V.: A lossless DEM compression for fast retrieval method using fuzzy clustering and MANFIS neural network. Eng. Appl. Artif. Intell. **29**, 33–42 (2014)
13. Martínez-Miranda, M.D., Nielsen, J.P., Wüthrich, M.V.: Statistical modelling and forecasting of outstanding liabilities in non-life insurance. SORT **36**(2), 195–218 (2012)
14. Wüthrich, M.V., Merz, M.: Stochastic Claims Reserving Methods in Insurance. John Wiley & Sons, West Sussex (2008)
15. Siddique, N., Adeli, H.: Computational Intelligence: Synergies of Fuzzy Logic, Neural Networks and Evolutionary Computing. John Wiley & Sons (2013)
16. England, P.D., Verrall, R.J.: Stochastic claims reserving in general insurance. Br. Act. J. **8**, 443–518 (2002)
17. Mulquiney, P., et al.: Artificial Neural Networks in Insurance Loss Reserving (2011)
18. Mack, T.: Measuring the Variability of Chain Ladder Reserve Estimates, Casualty Actuarial Society Forum, pp. 101–182 (1994)
19. Kaymak, U.: Defining a financial forecasting model for healthcare insurance companies, Eindhoven University of Technology (2013)
20. UCI Machine Learning Repository, Insurance Company Benchmark (COIL 2000) Data Set (2015). https://archive.ics.uci.edu/ml/datasets/Insurance+Company+Benchmark+(COIL+2000)
21. Beaver, W.H., McNichols, M.F., Nelson, K.K.: Management of the loss reserve accrual and the distribution of earnings in the property-casualty insurance industry. J. Acc. Econ. **35**, 347–376 (2003)
22. Zhang, Y., Dukic, V., Guszcza, J.: A Bayesian non-linear model for forecasting insurance loss payments. J. R. Stat. Soc. Ser. A (Stat. Soc.) **175**(2), 637–656 (2012)

A New Schema to Identify S-farnesyl Cysteine Prenylation Sites with Substrate Motifs

Van-Nui Nguyen[1(✉)], Thi-Xuan Tran[2], Hai-Minh Nguyen[1], Hong-Tan Nguyen[1], and Tzong-Yi Lee[3,4(✉)]

[1] University of Information and Communication Technology, Thai Nguyen, Vietnam
{nvnui,nhminh,nhtan}@ictu.edu.vn
[2] University of Economics and Business Administration, Thai Nguyen, Vietnam
tranxuantbhd84@gmail.com
[3] Department of Computer Science and Engineering, Yuan Ze University, Taoyuan City, Taiwan
francis@saturn.yzu.edu.tw
[4] Innovation Center for Big Data and Digital Convergence, Yuan Ze University, Taoyuan City, Taiwan

Abstract. Protein prenylation is the addition of hydrophobic molecules to a protein or chemical compound. It is a post-translational modification that plays very important roles for many cellular processes such as DNA replication, signaling, trafficking, and other cellular functions in eukaryotes. Protein S-farnesyl cysteine prenylation is a specific kind of prenylation involved in the transfer of a farnesyl moiety to a cytoplasmic cysteine at or near the C-terminus of the target protein. Recent advancements in proteomic technology have stimulated an increasing interested in the identification of protein S-farnesyl cysteine prenylation sites. However, there is still a lack of methods proposed for the prediction of S-farnesyl cysteine sites. With a rapidly increasing number of experimentally verified S-farnesyl cysteine sites, it is motivated in proposed new method for identifying S-farnesyl cysteine prenylation sites.

Keywords: Protein prenylation · Protein S-farnesyl cysteine prenylation · Substrate motif · Maximal Dependence Decomposition · Support vector machine

1 Introduction

Protein prenylation (also known as isoprenylation or lipidation), which was first discovered in fungi in 1978 [1], was the addition of hydrophobic molecules to a protein or chemical compound. The first prenylated protein in mammalian cells, farnesylated lamin B, was detected about ten year later [2,3]. Protein prenylation assumed that prenyl groups (3-methyl-but-2-en-1-yl) facilitate attachment to cell membranes, similar to lipid anchors like GPI anchor, though direct evidence was missing. Prenyl groups have been shown to be important for protein–protein binding through specialized prenyl-binding domains. In eukaryote, protein prenylation was a post-translational modification critical for many cellular

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_10

processes such as DNA replication, signaling and trafficking. It was mediated by protein farnesyltransferase (PFT) by recognizing 'CAAX' motif on protein substrate [4]. The process of protein prenylation was facilitated by three eukaryotic enzymes with partially overlapping substrate specificities: farnesyl transferase, Caax protease and geranylgeranyl transferase [5]. Protein S-farnesyl cysteine prenylation involved the transfer of a farnesyl moiety to a cytoplasmic cysteine at or near the C-terminus of the target protein. Farnesyltransferase (FT) recognized the so-called C-terminal CaaX box of substrate proteins to attach a farnesyl (15 carbons) anchor to the conserved cysteine via a thioether linkage [6].

Due to the very important role caused by protein S-farnesyl cysteine (SFC) prenylation, the amount of interests in the characterization of S-farnesyl cysteine prenylation has been increasing rapidly recently [5,7–12]. For example, Palsuledesai et al. [7] exhibited protein prenylation as an important posttranslational modification responsible for the correct cellular localization, activity, and protein-protein interactions of a number of signaling proteins. The authors have also indicated that prenyltransferase inhibitors could be used to suppress the activity of oncogenic Ras proteins to achieve antitumor activity, and played roles as potential therapeutics against several diseases including cancers, progeria, aging, parasitic diseases, and bacterial and viral infections.

Recently, several predictors have been designed to aid researchers in the prediction of S-farnesyl cysteine prenylation sites [6,13,14]. PRENbase [6] and PrePS [14] predictors were constructed on the basis of sequence and physical property profiles, including interpositional correlations, to characterize and predict partially overlapping substrate specificities, which is of medical importance in the case of understanding cellular action of FT inhibitors as anticancer and anti-parasite agents. In a recent work, Yubin Xie et al. [13] developed a robust tool named GPS-Lipid for predicting multiple lipid modification sites, including S-farnesyl cysteine prenylation sites [13]. The Swarm Optimization with an aging leader and challengers (ALC-PSO) [15] was also applied to obtain higher performance for predictor of GPS-Lipid. However, at the moment, there is a lack of computation models or tools for characterization as well as prediction of protein S-farnesyl cysteine prenylation sites. Furthermore, as more and more experimentally verified S-farnesyl cysteine prenylation sites become available, the lack of model for identification of S-farnesyl cysteine is serious.

Continue with previous works [16–19], we were motivated to propose a novel scheme for identifying S-farnesyl cysteine prenylation sites. Various features have been assessed in this work. The results revealed that our proposed model constructed based on hybrid feature of "AAC+AAPC+PSSM" could yield the best performance. Evaluation by five-fold cross-validation on the training data, the model generated an accuracy of 94.14 % and MCC of 0.850. When tested using independent testing, the model yielded an accuracy of 95.00 % and MCC of 0.747. In addition, the Maximal Dependence Decomposition (MDD) has been exploited to identify S-farnesyl cysteine prenylation sites with substrate motif. The eight MDDLogo-identified substrate motifs have been shown to provide a meaningful help for researchers in the characterization of S-farnesyl cysteine prenylation sites.

2 Material and Methods

2.1 Data Collection and Pre-processing

Experimentally verified S-farnesyl cysteine prenylation sites were collected from open resources and published literatures, including 711 proteins from UniProt/Swiss-Prot [20] (date: May, 2016), 117 proteins from dbPTM3.0 [21], 113 proteins from PRENbase [6], 97 proteins from GPS-Lipid [13], and 27 proteins from HPRD9.0 [22]. Details of these datasets were displayed in Table 1. After some technical steps to remove duplicate or redundant proteins, we obtained the final non-redundant dataset containing 670 uniques proteins with 718 S-farnesyl cysteine prenylation sites (positive data). To prepare for independence testing, we randomly selected 70 proteins from the non-redundant dataset to be as independent testing dataset. The remaining data was considered as training dataset. As a result, in this work, our final training dataset contained 600 uniques proteins, and the final independent testing dataset contained 70 uniques proteins.

Table 1. Data statistics of experimentally verified S-farnesyl cysteine prenylation sites collected from various open resources

Resources	S-farnesylated proteins	S-farnesyl cysteine sites	non-S-farnesyl cysteine sites
UniProt-5.2016	711	735	-
dbPTM	117	169	-
PRENbase	113	113	-
GPS-Lipid	97	106	-
HPRD 9.0	27	39	-
Total	1065	1162	-
Combined non-redundant dataset	670	718	-
Training dataset	600	634	5808
Independent testing dataset	70	84	954

In this work, we focused on the sequence-based characterization of S-farnesyl cysteine sites with substrate motif. So, the window length of 2n + 1 was utilized to extract sequence fragments centering at the experimentally verified ubiquitinated S-farnesyl cysteine (C) residue as well as containing n upstream and n downstream flanking amino acids. Given 600 experimentally verified ubiquitinated proteins, the sequence fragments containing window length of 2n + 1 amino acids and centering at lysine residue without the annotation of

S-farnesyl cysteine prenylation were regarded as the negative training data (non-ubiquitinated sites). According to a previous work [16,17] and our preliminary evaluation by using various window lengths, the window size of 13 (n = 6) has been shown to provide the optimal accuracy in the identification of S-farnesyl cysteine prenylation sites. By using a window size of 13, consequently, the training dataset contained 634 positive training sequences and 5808 negative training sequences. Due to the fact that some negative data (non-SFC fragment) were possible identical with positive data (SFC fragment) in the training dataset, as well as cross the testing dataset, the performance of the predictive model may be overestimated. Therefore, to avoid the overestimation of the model, it was necessary to apply CD-HIT program [23] to remove homologous data. After filtered out the homologous data with 40 % sequence identity, the training dataset consisted of 296 positive training sequences and 1051 negative training sequences.

The independent testing dataset was randomly selected from the non-redundant dataset, containing 70 proteins. The positive and negative independent testing dataset were constructed using the same approach as applied to the training dataset. Besides, the program cd-hit-2d, using sequence identity cutoff at 100 %, was applied again to remove the data redundancy between independent testing dataset and training dataset. This resulted in the final independent testing dataset containing 28 positive and 332 negative data.

2.2 Features Extraction and Encoding

In order to construct the predictive models for the identification of S-farnesyl cysteine prenylation sites, support vector machine was adopted to distinguish S-farnesyl from non-S-farnesyl cysteine sites based on sequence-based features, including: Amino Acid Composition (AAC), Amino Acid Pairwise Composition, and Evolutionary information (PSSM, Position-Specific Scoring Matrix) [16,17,24]. Besides, the four hybrid features was formed by combining single features being used for the investigation, including: "AAC+AAPC", "AAC+PSSM", "AAPC+PSSM", and "AAC+AAPC+PSSM".

2.3 Model Construction, Learning and Evaluation

Support vector machine (SVM) was adopted to construct the predictive models, and then learn the SVM classifiers based on extracted features. According to binary classification, the SVM using a kernel function maps the input samples into a higher dimensional space, and then finds out a a hyper-plane to discriminate between the two classes with maximal margin and minimal error. In this work, a public SVM library, LibSVM [25], was utilized to implement the predictive models for discriminating the S-farnesyl cysteine prenylation sites from non-S-farnesyl cysteine prenylation sites. The radial basis function (RBF) was selected as the kernel function for learning in the SVM classifiers, defined follows the formular: $K(S_i, S_j) = \exp\left(-\gamma \|S_i - S_j\|^2\right)$. In the SVM learning, two supporting factors to enhance the performance of the models are cost and gamma.

The RBF kernel function was determined by the gamma value, whereas the hyper-plane softness was controlled by cost value. To find the best final model, the predictive performance of models using different parameters was evaluated by performing five-fold cross-validation. The five-fold cross-validation has advantages in improving the reliability of evaluation because it considers all original data are regarded as both training and testing dataset, with each data is used for validation exactly once [16–19,26]. In order to assess the predictive performance of trained models, the followings measures were often used: Sensitivity (SEN), Specificity (SPE), Accuracy (ACC), and Matthews Correlation Coefficient (MCC): $SEN = TP/(TP+FN)$; $SPE = TN/(TN+FP)$; $ACC = TP/(TP+FN)$; $MCC = (TP \times TN)-(FN \times FP)TP+FN \times TN+FPTP+FPTN+FN$ Where: TP, TN, FP and FN represent the numbers of true positives, true negatives, false positives and false negatives, respectively.

Sensitivity indicated the rate of correct prediction on positive data (S-farnesylated cysteines), whereas specificity showed that on negative data (non-S-farnesylated cysteines). Accuracy pointed out the overall proportion of correctly predicted on both positive data and negative data. For the binary classification, accuracy was sometimes not useful in case of the two classes having very different sizes. Hence, MCC was a good choice and often used as a meaningful balanced measurement, even if sizes of two classes were very different [27]. The MCC value ranged from −1 to +1, whereas the three other measures (SEN, SPE, and ACC) ranged from 0 to 1. The coefficient value of +1 indicated that the prediction was perfect. In contrast, the coefficient values of 0 and −1 represent random and opposite predictions, respectively. Overall, a higher positive MCC value indicates a better prediction for correctly classifying positive and negative data.

2.4 Substrate Motif Discovery for the Identification of S-farnesyl Cysteine Sites

Recent advancements of bio-technology and informatics on high-throughput of mass strometry-based proteomics, make a rapid increasing number of experimentally verified S-farnesyl cysteine prenylation being available for researchers. However, the there still a lack of clue to help identify the S-farnesyl cysteine sites. Therefore, we were motivated to discover the potential substrate motif of S-farnesyl cysteine sites. In this work, Maximal Dependence Decomposition (MDD) [18,28] was adopted to explore substrate motif for the identification of S-farnesyl cysteine sites. MDD was shown to be effective in clustering splice sites for the purpose of splice site prediction, as well as identifying useful substrate motifs [16–18,28]. MDD adopts the chi-square test to assess the dependence of amino acid occurrence between two positions Ai and Aj that surround the S-farnesyl cysteine prenylation.

3 Results and Discussion

3.1 Impact of Single Features in Identifying S-farnesyl Cysteine Sites

To examine the position-specific amino acid composition for S-farnesyl cysteine prenylation sites, WebLogo [29] was applied to generate the graphical sequence logo for the relative frequency of the corresponding amino acid at positions surrounding S-farnesyl cysteine sites (at position 0). The identified motifs were subsequently evaluated on their ability to distinguish S-farnesyl cysteine prenylation from non-S-farnesyl cysteine prenylation by five-fold cross-validation.

Investigation of the differences between the AAC surrounding S-farnesyl cysteine prenylation and those of non-S-farnesyl cysteine prenylation showed that the overall trends were similar with slight variations. As shown in Fig. 1 (c), prominent amino acid residues included Ala (A), Ser (S), Gly (G), and Lys (K), and Met (M); while Trp (W), Try (Y), and Phe (F), were three of the least significant amino acid residues. Sequence logo displayed the most enriched residues surrounding the S-farnesyl cysteine prenylation (Cysteine C). As shown in Fig. 1(a), it also showed that the most conserved amino acid residues including of Phe (F), Lys (K), Ser (S), Met (M), and Val (V). In addition, the difference between SFC-sites and non-SFC sites was visualized using TwoSampleLogo [30]. The enriched residues appeared to be Phe (F), Pro (P), Ser (S), Gly (G) and Met (M); whereas the depleted amino acid residues included Val (V), Leu (L), Glu (E), Lys (K), and Gly (D) (Fig. 1(b)). An SVM model was trained to examine the effectiveness of AAC in identifying S-farnesyl cysteine prenylation. This SVM model used a 20-dimensional vector comprising of the composition scores for twenty types of amino acids. In order to evaluate the AAC-based model, the five-fold cross-validation was applied. As shown in Table 2, the model yielded 91.91 % accuracy, and an MCC value of 0.7998. Also, the AAPC-based model was trained to investigate the ability of AAPC and PSSM in identifying S-farnesyl cysteine prenylation. The accuracy and MCC of the AAPC-based model reached 88.27 % and 73.78, respectively. In addition to the composition of flanking amino acids, the evolutionary information (PSSM) was also investigated. As presented in Table 2, the five-fold cross-validation showed that the PSSM-based models yielded 92.68 % accuracy, and the MCC value of 0.807.

3.2 Impact of Hybrid Features in Identifying S-farnesyl Cysteine Sites

It is straightforward and very beneficial to combine two or more different approaches in machine learning to exploit advantages from them. Various methods have been applied to predict protein sites [16–18,31]. In our approach, hybrid features were built from the incorporation of two or more single features in order to form new features for the investigation. As a consequence, the hybrid features

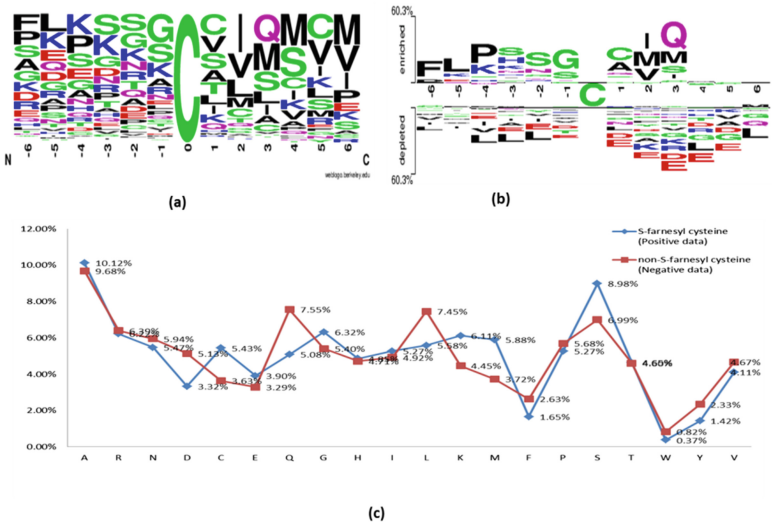


Fig. 1. The graphical sequence logo showing the relative frequency of the corresponding amino acid at positions surrounding S-farnesyl cysteine sites

Table 2. Performance evaluation by five-fold cross-validation

Feature	TP	FP	TN	FN	SEN	SPE	ACC	MCC
AAC	286	100	952	9	96.95 %	90.49 %	91.91 %	0.800
AAPC	291	153	898	5	98.31 %	85.44 %	88.27 %	0.738
PSSM	285	95	1058	11	96.28 %	91.76 %	92.68 %	0.807
AAC+AAPC	289	74	977	10	96.66 %	92.96 %	93.78 %	0.839
AAC+PSSM	286	67	983	14	95.33 %	93.62 %	94.00 %	0.842
AAPC+PSSM	286	68	982	14	95.33 %	93.52 %	93.93 %	0.840
AAC+AAPC+PSSM	**291**	**74**	**977**	**5**	**98.31 %**	**92.96 %**	**94.14 %**	**0.850**

were found to be the most effective in predicting protein S-farnesyl cysteine prenylation sites.

The performance of the model when tested with the hybrid features using the training data and independent testing data was shown in Tables 2 and 3, respectively. The hybrid feature "AAC+AAPC+PSSM" has been demonstrated to generate the best model which achieved the highest performance, with 94.14 % accuracy, and an MCC value of 0.8503. This indicates that the hybrid feature "AAC+AAPC+PSSM" would generate the most promising prediction results.

3.3 Independent Testing Performance

As mentioned previously, to assess the practicability of the trained models, an independent testing data set was constructed by randomly selected 70 uniques

Table 3. Performance evaluation by independent testing

Feature	TP	FP	TN	FN	SEN	SPE	ACC	MCC
AAC	24	25	307	4	85.71 %	92.47 %	91.94 %	0.611
AAPC	25	20	312	3	89.29 %	93.98 %	93.61 %	0.674
PSSM	25	19	313	3	89.29 %	94.28 %	93.89 %	0.683
AAC+AAPC	26	18	314	2	92.86 %	94.58 %	94.44 %	0.715
AAC+PSSM	25	19	313	3	89.29 %	94.28 %	93.89 %	0.683
AAPC+PSSM	24	19	313	4	85.71 %	94.28 %	93.61 %	0.661
AAC+AAPC+PSSM	**27**	**17**	**315**	**1**	**96.43 %**	**94.88 %**	**95.00 %**	**0.747**

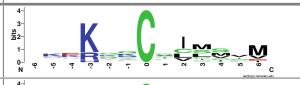
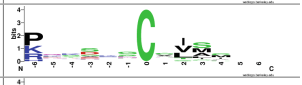
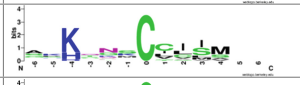
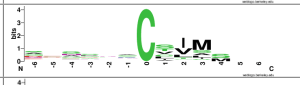
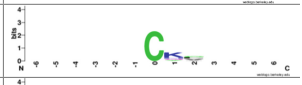
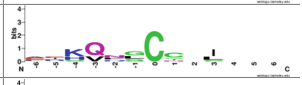
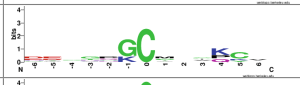
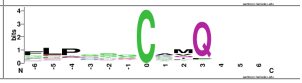
proteins from the final-non-redundant data. After several technical steps and data pre-processing, the independent testing data set comprised 28 positive and 332 negative data. The performance of the model when tested on the independent testing data set was shown in Table 3. The model constructed with the hybrid feature "AAC+AAPC+PSSM" delivered the best performance, with 95.00 % accuracy, and an MCC value of 0.747. This evidenced for the strength of our proposed method. Furthermore, this suggested that the hybrid approach of combining single features could be an effective and promising approach.

3.4 Substrate Motif Discovery for the Identification of S-farnesyl Cysteine Sites

MDD adopted a recursive chi-square test to evaluate the dependence of amino acid occurrence between two position surround the SFC-sites. In this work, MDD was applied to sub-divide the positive training data (296 SFC-site fragments) to eight subgroups containing significant substrate motifs. The negative data for each MDD-clustered subgroups were randomly selected from the negative training (1051 non-SFC-site fragments) with a ratio approximately equal to 1:3.551 (same as the ratio of positive training to negative training–296:1051). As a result, the eight useful substrate motifs were displayed in Table 4.

In addition, MDD-clusters containing Lysine (K), Proline (P) and Glutamine (Q) residues in conserved motifs appeared to generate better performances. For example, evaluation by independent testing revealed that the MDD cluster 2, consisting of Proline (P) and Lysine (K) residues at position −6 of conserved motifs, yielded an accuracy of 98.61 %. Similarly, MDD cluster 3, which was comprised of Lysine (K) residues at position −4 in conserved motifs, obtained 98.19 % accuracy. In general, almost all clusters containing conserved Lysine (K), Proline (P) and Glutamine (Q) at specific positions, could yield good performance. In contrast, other clusters without clearly conserved motifs containing Lysine (K), Proline (P) and Glutamine (Q) residues generally showed lower sensitivity. This suggests that, for S-farnesyl cysteine prenylation, the substrate site specificities may depend on the conserved position of Lysine (K), Proline (P) and Glutamine (Q) residues.

Table 4. Substrate motif detected by MDD

MDD-clustered subgroup	Number of positive data	Number of negative data	ACC	MCC	Substrate site motif
1	62	220	98.33 %	0.893	
2	53	188	98.61 %	0.914	
3	19	67	98.33 %	0.899	
4	65	231	91.67 %	0.557	
5	7	25	91.94 %	0.580	
6	7	25	95.83 %	0.766	
7	8	28	96.11 %	0.788	
8	75	267	98.06 %	0.885	

4 Conclusion

Protein S-farnesyl cysteine prenylation was a kind of post-translational modification that plays critical roles for many cellular processes such as DNA replication, signaling, and trafficking, found in all eukaryotic cells. It comprised an attachment of S-farnesyl isoprenoid, which are typically involved in mediating not only protein-membrane but also protein-protein interactions. Inhibition of S-farnesyl cysteine prenylation has been extensively investigated to suppress the activity of oncogenic Ras protein to achieve antitumor activity. The current status of prenyltransferase inhibitors have been accounted to be as potentially therapeutics against several diseases, including: cancers, progeria, aging, parasitic diseases, bacterial and viral infections. In this study, we proposed a new schema to identify protein S-farnesyl cysteine prenylation sites with substrate motifs by exploiting maximal dependence decomposition method. The SVM models based on various features have been constructed and investigated. The hybrid feature "AAC+AAPC+PSSM" has been found to generate the best model that yields the highest performance. Evaluation of the proposed model using an independent testing revealed the strength of our proposed method in comparison with existing prediction tools. In addition, the eight useful substrate motifs were discovered proving promising clue for biologist to recognize the protein S-farnesyl cysteine prenylation sites.

References

1. Kamiya, Y., et al.: Structure of rhodotorucine A, a novel lipopeptide, inducing mating tube formation in Rhodosporidium toruloides. Biochem. Biophys. Res. Commun. **83**(3), 1077–1083 (1978)
2. Farnsworth, C.C., et al.: Human lamin B contains a farnesylated cysteine residue. J. Biol. Chem. **264**(34), 20422–20429 (1989)
3. Wolda, S.L., Glomset, J.A.: Evidence for modification of lamin B by a product of mevalonic acid. J. Biol. Chem. **263**(13), 5997–6000 (1988)
4. Soni, R., et al.: Structure-based binding between protein farnesyl transferase and PRL-PTP of malaria parasite: an interaction study of prenylation process in Plasmodium. J. Biomol. Struct. Dyn., 1–12 (2016)
5. Novelli, G., D'Apice, M.R.: Protein farnesylation and disease. J. Inherit. Metab. Dis. **35**(5), 917–926 (2012)
6. Maurer-Stroh, S., et al.: Towards complete sets of farnesylated and geranylgeranylated proteins. PLoS Comput. Biol. **3**(4), e66 (2007)
7. Palsuledesai, C.C., Distefano, M.D.: Protein prenylation: enzymes, therapeutics, and biotechnology applications. ACS Chem. Biol. **10**(1), 51–62 (2015)
8. Hechinger, A.K., et al.: Inhibition of protein geranylgeranylation and farnesylation protects against graft-versus-host disease via effects on CD4 effector T cells. Haematologica **98**(1), 31–40 (2013)
9. Charron, G., et al.: Prenylome profiling reveals S-farnesylation is crucial for membrane targeting and antiviral activity of ZAP long-isoform. Proc. Natl. Acad. Sci. USA **110**(27), 11085–11090 (2013)
10. Geryk-Hall, M., Yang, Y., Hughes, D.P.: Driven to death: inhibition of farnesylation increases Ras activity and promotes growth arrest and cell death [corrected]. Mol. Cancer Ther. **9**(5), 1111–1119 (2010)
11. Goodsell, D.S.: The molecular perspective: protein farnesyltransferase. Oncologist **8**(6), 597–598 (2003)
12. Einav, S., Glenn, J.S.: Prenylation inhibitors: a novel class of antiviral agents. J. Antimicrob. Chemother. **52**(6), 883–886 (2003)
13. Xie, Y., et al.: GPS-Lipid: a robust tool for the prediction of multiple lipid modification sites, Scientific reports, 6, 28249 (2016)
14. Maurer-Stroh, S., Eisenhaber, F.: Refinement and prediction of protein prenylation motifs. Genome Biol. **6**(6), R55 (2005)
15. Chen, W.N., et al.: Particle swarm optimization with an aging leader and challengers. IEEE Trans. Evol. Comput. **17**(2), 241–258 (2013)
16. Nguyen, V.N., et al.: A new scheme to characterize and identify protein ubiquitination sites. In: IEEE/ACM transactions on computational biology and bioinformatics IEEE, ACM (2016)
17. Nguyen, V.N., et al.: Characterization and identification of ubiquitin conjugation sites with E3 ligase recognition specificities. BMC Bioinform. **16**(Suppl. 1), S1 (2015)
18. Lee, T.Y., et al.: Exploiting maximal dependence decomposition to identify conserved motifs from a group of aligned signal sequences. Bioinformatics **27**(13), 1780–1787 (2011)
19. Lee, T.Y., et al.: SNOSite: exploiting maximal dependence decomposition to identify cysteine S-nitrosylation with substrate site specificity. PLoS One **6**(7), e21849 (2011)

20. Boeckmann, B., et al.: The SWISS-PROT protein knowledgebase and its supplement TrEMBL in 2003. Nucleic Acids Res. **31**(1), 365–370 (2003)
21. Lu, C.T., et al.: DbPTM 3.0: an informative resource for investigating substrate site specificity and functional association of protein post-translational modifications. Nucleic Acids Res. **41**(Database issue), D295–D305 (2013)
22. Keshava Prasad, T.S., et al.: Human protein reference database–2009 update. Nucleic Acids Res. **37**(Database issue), D767–D772 (2009)
23. Huang, Y., et al.: CD-HIT suite: a web server for clustering and comparing biological sequences. Bioinformatics **26**(5), 680–682 (2010)
24. Altschul, S.F., et al.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**(17), 3389–3402 (1997)
25. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2** (2011)
26. Nguyen, V.N., et al.: UbiNet: an online resource for exploring the functional associations and regulatory networks of protein ubiquitylation. Database : J. Biol. Databases and Curation (2016)
27. Matthews, B.W.: Comparison of the predicted and observed secondary structure of T4 phage lysozyme. Biochim. Biophys. Acta **405**(2), 442–451 (1975)
28. Burge, C., Karlin, S.: Prediction of complete gene structures in human genomic DNA. J. Mol. Biol. **268**(1), 78–94 (1997)
29. Crooks, G.E., et al.: WebLogo: a sequence logo generator. Genome Res. **14**(6), 1188–1190 (2004)
30. Vacic, V., Iakoucheva, L.M., Radivojac, P.: Two sample logo: a graphical representation of the differences between two sets of sequence alignments. Bioinformatics **22**(12), 1536–1537 (2006)
31. Tung, C.W., Ho, S.Y.: Computational identification of ubiquitylation sites from protein sequences. BMC Bioinform. **9**, 310 (2008)

A Novel Framework Based on Deep Learning and Unmanned Aerial Vehicles to Assess the Quality of Rice Fields

Nguyen Cao Tri[1], Tran Van Hoai[1], Hieu N. Duong[1(✉)], Nguyen Thanh Trong[1], Vo Van Vinh[2], and Vaclav Snasel[3]

[1] Faculty of Computer Science and Engineering, Ho Chi Minh City University of Technology, Ho Chi Minh City, Vietnam
dnhieu@hcmut.edu.vn
[2] Tay Ninh Agricultural Breeding Center, Tay Ninh, Vietnam
[3] Faculty of Electrical Engineering and Computer Science, VSB-Technical University of Ostrava, Ostrava, Czech Republic

Abstract. In the past few decades, boosting crop yield has been extensively regarded in many agricultural countries, especially Vietnam. Due to food demands and impossibility of crop-field area increasing, precision farming is essential to improve agricultural production and productivity. In this paper, we propose a novel framework based on some advanced techniques including deep learning, unmanned aerial vehicles (UAVs) to assess the quality of Vietnamese rice fields. UAVs are responsible for taking images of the rice fields at low or very low altitudes. Then, these images with high resolution will be processed by the deep neural networks on high performance computing systems. The main task of deep neural networks is to classify the images into many classes corresponding to low and high qualities of the rice fields. To conduct experimental results, the rice fields located in Tay Ninh province are chosen as a case study. The experimental results indicate that this approach is quite appropriate for agricultural Vietnamese practice since its accuracy is approximately 0.72.

Keywords: Deep learning · Unmanned aerial vehicles · Precision agriculture · Rice fields

1 Introduction

Boosting crop yield has been regarded extensively in the past few decades. The yield trends are insufficient since the global population uncontrollably increases whereas crop-field areas significantly decrease. Global crop yield will need to double to meet the demands of 10 billions people by 2050 [10]. Due to impossibility of crop-field area increasing, precision farming is essential to improve agricultural production and productivity. Precision farming or precision agriculture (PA) is an interesting concept that applies advanced techniques, such as remote sensing, geographic information systems, telecommunication, high computational systems, etc. to observe, measure and respond to crop-field variability.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_11

Among several alternative techniques used to tackle the tasks of PA, applications of image analysis have attracted a huge amount of researchers.

Recently, there have been many interesting research topics in PA, such as detection of plant leaf diseases, crop-yield estimation, weed detection, fruit grading, and so on [11,13]. When dealing with a specific problem, researchers often base on their experience and experiments to select appropriate methods. Most commonly-used methods are combinations of image preprocessing techniques and machine learning algorithms [5,11]. While image preprocessing consists of some well-known methods, such as point processing, contrast stretching, histogram processing, mask processing, and so on [3], the machine learning algorithms consist of a huge amount of methods, such as k-nearest neighbors, neural networks, support vector machines, self-organizing map, fuzzy logic, etc. [5]. Among several machine learning algorithms, artificial neural networks (ANNs) have attracted plenty of researchers.

In [7], Ainong Li *et al.* utilized ANNs that were trained by the SCE-UA optimization algorithm to estimate crop yield. The authors also compared the performance of ANNs with a multivariate linear regression model (MLR). As a result, they indicated the proposed ANN model was superior to MRL since the prediction accuracy of ANNs was approximately 85 %. In 2003, Matthew J. Aitkenhead *et al.* used a self-organizing neural network model based on common neural networks and Genetic Algorithm to discriminate weed and crop [1]. The authors concluded that the neural network-based approach outperformed a simple morphological characteristic measurement of leaf shape. In [2], Heba Al-Hiary *et al.* proposed a hybrid approach based on k-means algorithm, skilful techniques of image processing and neural networks. The authors applied the approach to classify plant diseases. Through experiments, they proved that the approach was significantly effective since the precision of the approach was between 83 % and 94 %.

In recent years, deep learning exploiting the super strength of high performance computing (HPC) systems is becoming an interesting approach for the problems involving image classification in terms of accuracy. Particularly, deep learning in neural networks, so-called deep neural networks (DNNs), have won numerous contests in pattern recognition and machine learning [12]. To model sophisticated and incomplete real-world data, DNNs contain multiple processing layers with complex structures, consequently bringing up a very big computational complexity. Hence, DNNs need to be supported by HPC systems to become a powerful and useful tool. Due to the strict requirement of hardware, DNNs have been rarely used in PA.

In Vietnam, agriculture is one of the major industries and contributes significantly to the national Gross Domestic Product (GDP). Despite the trending away from agriculture, the Vietnamese economy has still contributed by the domestic agriculture with an approximate factor of 15−20 % in the last few years [9]. Moreover, Vietnam is among the top 5 rice export countries in the world, contributing about 7.4 % (equivalent to 1.8 billion USD), in 2015[1].

[1] http://www.worldstopexports.com/rice-exports-country.

Thus, it is necessary to significantly support Vietnamese agriculture in many aspects, such as national policies, advanced agriculture technologies, applications of computer science, and so on. To date, there has been scant significant research on computer science applications to support Vietnamese agriculture due to the vast distance between theory and practice.

Considering this urgent practical demand, in this study, we focus on applying advanced techniques including DNNs, unmanned aerial vehicles (UAVs) to assess the quality of Vietnamese rice fields. The quality of rice fields is defined as their density including sparse density and normal density. Note that the density of rice is an important factor that decides a successful harvest. We utilize UAVs to collect high-resolution imagery of the rice fields from low altitudes. Then, we use DNNs to classify the imagery to assess the quality of the rice fields. We conduct experiments on a HPC system that consists of several super computers. For image collecting, we choose the 10-day-old rice fields located in Tay Ninh province as a case study.

The rest of this paper is organized as follows. Section 2 presents a conceptual framework based on DNNs, UAVs. In Sect. 3, we present experimental results. Some finding also will be indicated in Sect. 3. Finally, we draw conclusion and perspectives for future work in Sect. 4.

2 Methodology

Figure 1 presents the conceptual framework that utilizes advanced techniques to assess the quality of rice fields. The framework consists of four main components as follows.

1. **Image acquisition.** UAVs are used to take images of rice fields at low altitude. The UAVs that are equipped with advanced hardware and software will directly send captured images to computers via a local wireless network.
2. **Image processing.** We apply several image processing techniques, such as brightness/contrast adjustment, segmentation, filtering, etc. to improve the quality of captured images and convert them to usable images as well.
3. **Image sampling.** We utilize the expertise and experience of agricultural experts to sample the preprocessed images. Then the sampled images are used by a deep learning-based classifier for classification objectives.
4. **Classifying the imagery by deep learning.** Deep neural networks (so-called deep learning-based classifiers) are used to learn the sampled imagery and then classify the rest imagery. Through the classification results of the deep learning-based classifiers, we provide significant assessments to farmers. To learn the deep learning-based classifiers, a high performance computing system is employed due to the big computational complexity of deep learning models.

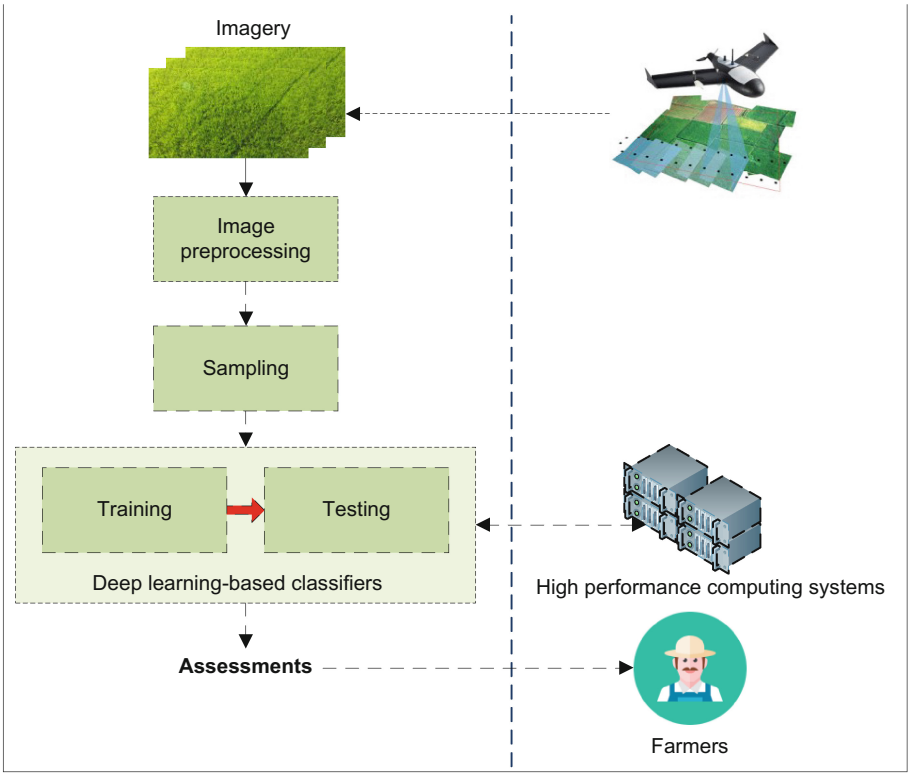


Fig. 1. The work flow of rice field assessments based on UAVs, deep learning.

2.1 Image Acquisition

The remote sensing and/or satellite imagery are common sources of most applications of image processing [5,11]. However, in this study, which is based on methods of deep learning and HPC systems, we need high-quality images as well as recent images of rice fields at some specific time and places.

The images also require to exactly cover the whole of rice fields that normally have small sizes and different shapes in Vietnam. According to the strict requirement, satellite images are unacceptable since their quality strongly depends on weather, cloud condition, low resolution, and low bit-rate. Moreover, the cost of satellite images is quite expensive and cannot be acquired at anytime as well.

In this study, Drone/UAVs are used to take images from low and very low altitude. The images taken from low attitude by Drone/UAVs give several significant features, such as low cost, high resolution, high quality/bit-rate, noise avoiding, etc. In addition, we can easily replace alternative camera types to capture multiple bands (Green, Near IR, Blue RED, RGB) to response to several different objectives.



Fig. 2. An illustration of APM flight controller.

To take images from rice fields, the UAVs, which are used in this study, are equipped with a flight controller, namely APM. The controller, as seen in Fig. 2, is responsible for several automatic pilot functions, such as automatically taking off, automatically setting waypoints, automatically landing, and so on. These features allow farmers setting waypoints to ensure that the taken images cover the whole of the rice fields according to altitudes of the UAV and camera's specifications.

2.2 Image Preprocessing

Image preprocessing is a very important step that significantly decides the results of any classifiers. The quality of images collected from rice fields by UAVs are impacted by some factors of surrounding environment, such as air, sunlight, wind, and so on. Hence, to improve the quality of images, the images are meticulously processed. Figure 3 presents fours steps of the image preprocessing.

- **Brightness/contrast adjusting.** Due to the influence of sun, the captured images are quite dark. Thus, it is necessary to brighten these images by adjusting their brightness and contrast.
- **Segmentation.** The original images captured by UAVs are quite large (4608×2592). This step segments these images into many smaller pieces whose sizes are 256×256.
- **Grayscale converting.** The task of this step is to convert color images to grayscale images. Although DNNs are able to work well with color images directly, the grayscale converting technique is also employed to speed up the computation of DNNs.

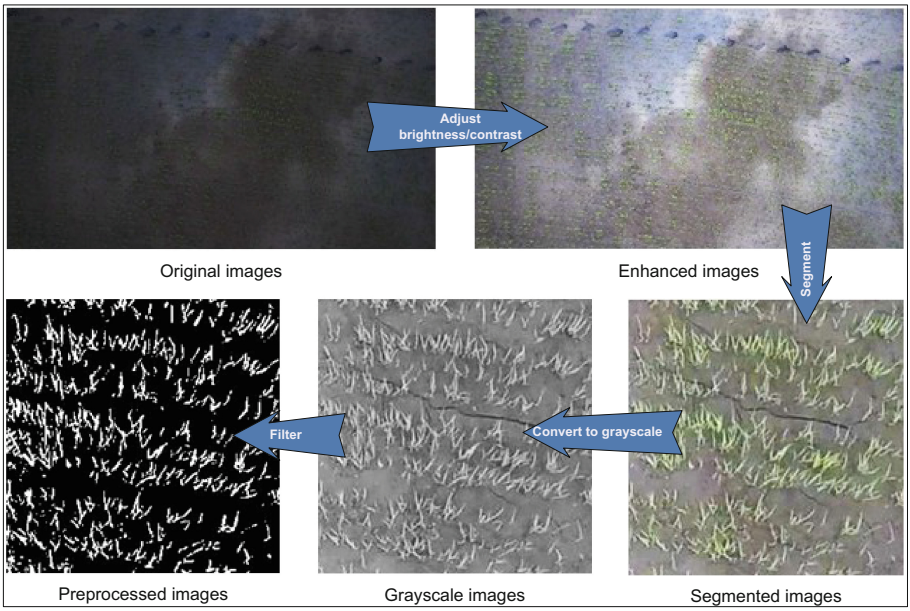


Fig. 3. The work flow of the image preprocessing.

– **Thresholding.** A threshold is used to separate two types of color objects: bright-gray objects corresponding to rice trees and the rest corresponding to soil, water, bubbles, etc.

2.3 Deep Neural Networks

Among several variants of ANNs in terms of deep learning, Convolutional Neural Networks (CNNs), which are used in this study, are ubiquitous and attracted a large amount of researchers. A CNN is a type of Feed-Forward Neural Network (FFNN) which are inspired by sophisticated biological processes [8]. Typically, a CNN consists of many layers that can be classified into two main types: convolution layer and pooling (sub-sampling) layer. That two types of layers are alternately arranged forms a massive structure [6]. The last layers are fully-connected layers that have full connections to the previous layers, as commonly seen in traditional ANNs. Figure 4 presents an example of CNN, namely LeNet-5. The interesting idea of CNNs can be shortly summarized by two key words: convolution and image resolution that correspond to convolution layers and pooling layers, respectively. Recently, deep learning in general and CNNs in particular have proved their power since they won many competitions of image classification [12]. This is the main reason why we choose this method to address the task of image classification.

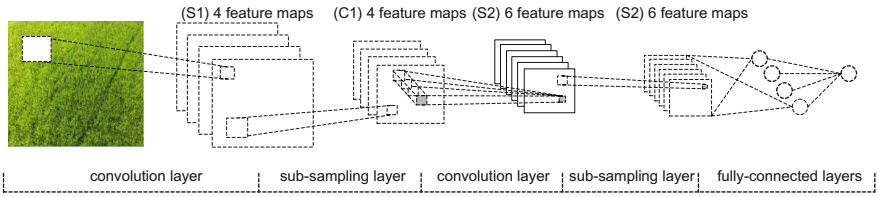


Fig. 4. A typical architecture of LeNet-5 [6].

3 Experimental Results

The imagery used in this study is collected at the rice fields located in Tay Ninh province, southern Vietnam. The particular task of this study is to assess the quality of 10-day-old rice fields. We collected the imagery in two rice fields whose the total area is approximately two hectares. The number of collected images is about 800. At the sampling step, based on expertise and experience of agricultural experts, we select 200 of preprocessed images and classify them into two classes: sparse density and normal density. For this dataset, we use 160 images for training and 40 images for testing CNNs.

Table 1. The setting of CNNs

Setting parameters of CNNs		
Names	Values	Notions
Convolution layers	3	- Receptive field: 5×5
		- Stride: 2×2
		- Zero-padding: 2×2
Feature maps	10, 20, 30	The number of feature maps of 1^{st}, 2^{nd}, and 3^{rd} convolution layers are 10, 20, 30, respectively
Pooling layers	3	- Receptive field: 5×5
		- Stride: 2×2
Nodes of fully-connected layer	100	
Output nodes	2	2 is the number of classes: sparse density and normal density
Learning rate	0.01	
Momentum	0.05	
Epoch	1000	
Iteration	1	
Training mode	Online	
Activation function	ReLU	

Table 2. Performance of CNNs

Training phases		
Names	Values	Formulas
MSE	0.301	$MSE = \frac{1}{n}\sum_{i=1}^{n}(O_i - P_i)^2$
Testing phases		
Names	Values	Formulas
Accuracy	0.720	$Accuracy = (TP+TN)/(TP+TN+FP+FN)$, where TP: True Positive, TN: True Negative, FP: False Positive, FN: False Negative
Precision	0.7619	$Precision = TP/(TP+FP)$
Recall	0.7115	$Recall = TP/(TP+FN)$
F1 Score	0.7359	$F1Score = 2*TP/(2TP+FP+FN)$

Actually, ones often confuse when setting parameters of CNNs, such as the number of convolution layers, the number of pooling layers, the number of feature maps, learning rates, momentum, etc. to adapt for a specific dataset. We apply the *trial-and-error* strategy to find out the most relevant setting of CNNs. Table 1 presents a setting of CNNs that gives a quite good result. Note that the deep model is executed on HPC system equipped compute nodes, namely XL250a GEN9, with 2 x E5-2680 v3, 128 GB RAM, and 2 x Intel Xeon Phi 7120P co-processors.

Table 2 presents the classification results of the deep learning-based classifier. The statistical indicators are quite acceptable since their values are approximately

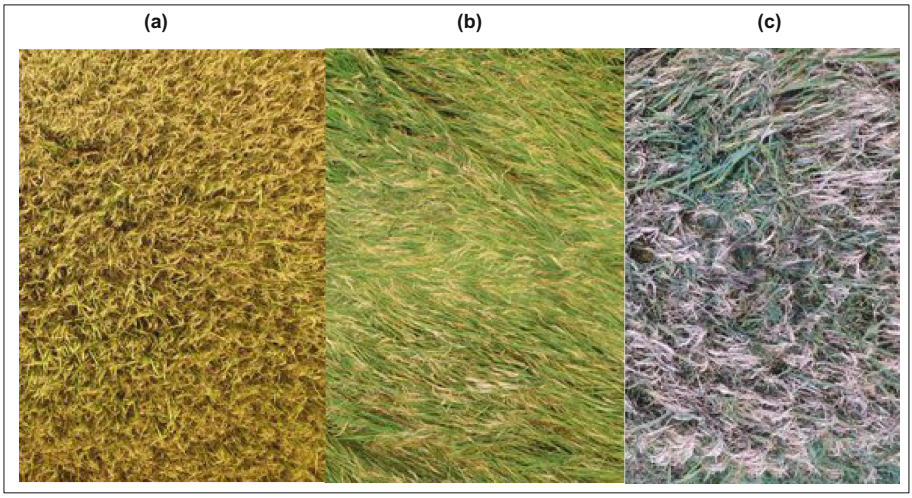


Fig. 5. A sample of imagery used for assessing rice-field productivity.

0.75. As mentioned above, the performance of CNNs depends on their setting parameters. The *trial-and-error* strategy is barely acceptable since the time consumption of DNNs is very big. For one setting, DNNs take about 48 hours to be convergent. Hence, some optimization algorithms may be taken into consideration in future work to explore the best setting of DNNs.

4 Conclusion

This paper introduced a novel framework based on some advanced techniques including deep learning, unmanned aerial vehicles to assess the quality of Vietnamese rice fields. To verify the performance of the framework, we use a dataset collected at the 10-day-old rice fields located in Tay Ninh province. The proposed framework is deployed to assess the quality of rice fields according to the criterion of density. The experimental results indicate that this approach is quite appropriate for agricultural Vietnamese practice since its accuracy is approximately 0.72. However, the performance of this approach can be improved. There are some drawbacks of this study that should be considered and tackled in future work as follows.

1. **The drawbacks of image processing.** The size of preprocessed images is quite large and thus they might contain lots of noise. Reducing these images can compress their useful information that can be easily learned by DNNs. Furthermore, with smaller-size images, learning time of DNNs will be significantly reduced. Besides, some other image processing techniques can be used to sharpen object borders.
2. **The drawbacks of DNNs.** Optimization algorithms should be used to find out the most relevant setting of parameters instead of the *trial-and-error* strategy. Moreover, we will run DNNs on compute nodes supporting Graphic Processing Unit (GPU) that can speed up learning algorithm several times, e.g. between 6.5× and 44× as introduced in [4].
3. **The size of the dataset.** DNNs need a big dataset to work effectively. Hence, some data augmentation techniques will be considered to increase the size of the dataset.

Moreover, in future research, the productivity of rice fields will be assessed by using this framework. The productivity will be assessed by analyzing images of 5-day-to-harvest rice fields as seen in Fig. 5. The figure illustrates three types of rice fields that are normality, wind-caused lying, disease infection corresponding to Fig. 5(a), (b) and (c), respectively.

Acknowledgements. The authors would like to thank Faculty of Computer Science and Engineering, HCMC University of Technology for providing facilities for this study. The applications presented in this paper were tested on the High Performance Computing Center (HPCC) of the faculty. This research was funded by HCMC Department of Science and Technology, under contract number 39/2015/HD-SKHCN.

References

1. Aitkenheada, M., et al.: Weed and crop discrimination using image analysis and artificial intelligence methods. Comput. Electron. Agric. **39**, 157–171 (2003)
2. Al-Hiary, H., et al.: Fast and accurate detection and classification of plant diseases. Int. J. Comput. Appl. **17**(1), 31–38 (2011)
3. Alginahi, Y.: Preprocessing techniques in character recognition. InTechOpen (2010)
4. Bergstra, J., et al.: Theano: a CPU and GPU math expression compiler. In: Proceedings of the Python for Scientific Computing Conference, pp. 1–7 (2010)
5. Ghaiwat, S.N., Arora, P.: Detection and classification of plant diseases using image processing techniques: a review. Comput. Electron. Agric. **2** (2014)
6. Lecun, Y., et al.: Gradient-based learning applied to document recognition. Proc. IEEE **86**, 2278–2324 (1998)
7. Li, A., Liang, S., Wang, A., Qin, J.: Estimating crop yield from multi-temporal satelite data using multivariate regression and neural network techniques. Photogram. Eng. Remote Sens. **73**(10), 1149–1157 (2007)
8. Matsugu, M., Mori, K., Mitari, Y., Kaneda, Y.: Subject independent facial expression recognition with robust face detection using a convolutional neural network. Neural Netw. **16**, 555–559 (2003)
9. McCaig, B., Pavcnik, N.: Moving out of agriculture: structural change in vietnam. The International Bureau of Economic Research (2013)
10. Ray, D.K., Mueller, N.D., West, P.C., Foley, J.A.: Yield trends are insufficient to double global crop production by 2050. Plos One **8**(6), e66428 (2013)
11. Sankaran, S., Misha, A., Ehsani, R., Davis, C.: A review of advanced techniques for detecting plant diseases. Comput. Electron. Agric. **72**, 1–13 (2010)
12. Schmidhuber, J.: Deep learning in neural networks: an overview. Neural netw. **61**, 85–117 (2015)
13. Vibhute, A., Bodhe, S.: Applications of image processing in agriculture: a survey. Int. J. Comput. Appl. **52**(2), 34–40 (2012)

A Semi-supervised Learning Method for Hybrid Filtering

Thi Lien Do(✉) and Duy Phuong Nguyen(✉)

Posts and Telecommunications Institue of Technology, Hanoi, Vietnam
{liendt,phuongnd}@ptit.edu.vn

Abstract. Recommender systems are the auto systems of providing appropriate information and removing unappropriate information for users. The recommender systems are built based on two main information filtering techniques: Collaborative filtering and content-based filtering. Content-based filtering performs effectively on documents representing as text but has problems to select information features on multimedia data. Collaborative filtering perform well on all types of information but had problems when sparse data, new uses and new items. In this paper, we propose a new unify model between collaborative filtering and content-based filtering by a semi-supervised learning method. The model is built based on two semi-supervised procedures: the first procedure semi-supervise ratings set between users and item's features, the second procedure semi-supervise ratings set between items and users features. The first procedure allows us to detect new items that is high suitable capability with the users. The second procedure allows us to detect new users that is high suitable ability with the items. Two procedures performed simultaneously and complement each other for suitable predicted values to improve recommender results. The experimental results on real data sets show that the proposed methods utilize effectively the advantages and limit significant disadvantages of baseline filtering methods.

Keywords: Collaborative filtering recommendation · Content-based filtering recommendation · Hybrid filtering recommendation system · Item-based recommendation · User-based recommendation

1 Introduction

Supposedly, we have a finite set U = $u_1, u_2, \ldots, u_N$ is the set of N users, P = p_1, p_2,.., p_M is the set of M items. Each item $p_x \in$ P can be paper, news, merchandise, movie, service or any informational types that the users need. Relationship between the users set U and the items set P are represented by the rating matrix R = $\{r_{ix}$: i = 1, 2, ..N; x = 1, 2,..M$\}$. Each value r_{ix} represents rating of the user $u_i \in$ U with the item $p_x \in$ P. Normally, r_{ix} having a value in the domain F = $\{1, 2,.., g\}$. The value r_{ix} can be collected directly by inquiring users opinion or indirectly by users feedback. The rating matrix R is the input matrix of collaborative filtering recommender systems [3,13]. For convenience

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_12

of presentation, I write short $p_x \in$ P as x∈P; $u_i \in$ U as i∈U. The letters i, j are always used to refer to users in next section of the paper. Each item x∈P is presented by $|C|$ content features, C = $\{c_1, c_1,.., c_{|C|}\}$. The content feature $c_s \in$ C can receive from feature selection methods in the field of information retrieval. For example x∈P is the movie then content features may represent for the movie are C=genre, producer, studio, actor, director.... Conventionally, w_x = $\{w_{x1}, w_{x2},.., w_{x|C|}\}$ is the weighted vector for content feature values of the item x∈P. Meanwhile, the weighted matrix W =$\{w_{xs}$: x =1, 2, .., M; s =1, 2, .., $|C|\}$ is the input of content-based recommender systems based on information of items [2,3,6]. For convenience in representation, I write short $c_s \in$ C as s∈C. The letter s is always used to refer to item content feature in next section of the paper. Each user i∈U is presented by $|T|$ content features, T = $\{t_1, t_2,..,$ $t_{|T|}\}$. The content feature $t_q \in$ T is usually individual information of each user (Demographic Information). For example, content features of the user i∈U can be T = gender,age, occupation, degree,.... Conventionally, $v_i = \{v_{i1}, v_{i2},.., v_{i|T|}\}$ is the weighted vector for content feature values of the user i∈U. Meanwhile, the weighted matrix V = $\{v_{iq}$: i =1, 2, .., N; q =1, 2, .., $|T|\}$ is the input of content-based recommender systems based on information of users [3,5]. For convenience in representation, I write short $t_q \in$ T as q∈T. The letter q is always used to refer to user content feature in next section of the paper. Next, we sign $P_i \in$ P is the item set x∈P that is rated by the user i∈U and $U_x \in$ U is the user set i∈U that gave rating to the item x∈P. With each user that need recommendation i∈U (known as the current user, the user need to be recommended or the active user), tasking of recommender methods is suggesting K items x∈(P$\setminus P_i$) that appropriate with the user i. In this paper, we propose a unify model between collaborative filtering and content-based filtering by a semi-supervised learning method to take advantages and limit difficulty of baseline filtering methods. To focus on the new propositions of the paper, in the next section, we present a method to estimate weights between user content features and items of content-based filtering. In the Sect. 3, we present the semi-supervised learning method based on user ratings, item features, user features. In the Sect. 4, we present a experimental method and compare proposed methods with baseline methods. In the last section is conclusion and developable direction of the paper.

2 A Unify Method to Represent Values of Content Features

No limitting generality of the problem are represented by Sect. 1, we assume rating value of the user i∈U with the item x∈P be determined by the formula (1). Each item x∈P is presented by $|C|$ content features, C=$\{c_1, c_1,.., c_{|C|}\}$ is determined by the formula (2). Each user i∈U is presented by $|T|$ content features, T=$\{t_1, t_2,.., t_{|T|}\}$ is determined by the formula (3).

$$r_{ix} = \begin{cases} v & If\ the\ user\ i\ rate\ the\ item\ x\ with\ v \\ & level\,(v \in F) \\ 0 & If\ the\ user\ i\ hasn't\ known\ the\ item\ x\ yet \end{cases} \quad (1)$$

$$c_{xs} = \begin{cases} 1 & If\ the\ item\ x\ has\ the\ content\ feature\ s \\ 0 & If\ the\ item\ x\ hasn't\ the\ content\ feature\ s \end{cases} \quad (2)$$

$$t_{iq} = \begin{cases} 1 & If\ the\ user\ i\ has\ the\ content\ feature\ q \\ 0 & If\ the\ user\ i\ hasn^{'}t the\ content\ feature\ q \end{cases} \quad (3)$$

2.1 Unifing User Profiles of Content-Based Filtering into the Rating Matrix

To build profiles using items for the user, we need performing 2 tasks: determining the set of items that the user had ever accessed or used in the past and estimating weight for each item content feature in user profiles [2,4,6]. Symbol $P_i \in$ P is determined by the formula (4) is the set of items that the user i∈U rated the item x∈P.

$$P_i = \{x \in P \mid r_{ix} \neq 0 \quad (i \in U, x \in P\ )\} \quad (4)$$

Symbol Item(i, s) is the set of items x $\in P_i$ containing the item content feature s∈C be determined by the formula (6).

$$Item\,(i,s) = \{x \in P_i \mid c_{xs} \neq 0 \quad (i \in U, x \in P,\ s \in C\ \} \quad (5)$$

Based on P_i and Item(i, s), content-based recommender methods estimate weight w_{is} reflecting importance of the item content feature s to the user i. If the value |Item(i, s)| exceeds a certain threshold θ then weight of the item content feature s∈C with the user i∈U is w_{is} that be calculated by average of all rating values. In the otherhand, if |Item(i, s)| is less than θ, the value w_{is} is calculated by sum of all rating values then divide for θ. Based on w_{is}, we extend the rating matrix following the formula (7)

$$w_{is} = \begin{cases} \frac{1}{|Item(i,s)|} \sum_{x \in Item(i,s)} r_{ix} & If\ |Item\,(i,x)| \geq \theta \\ \frac{1}{\theta} \sum_{x \in Item(i,s)} r_{ix} & If\ \ |Item\,(i,x)| < \theta \end{cases} \quad (6)$$

$$r_{ix} = \begin{cases} r_{ix} & If\ x \in P \\ w_{is} & If\ s \in C\ (x = s) \end{cases} \quad (7)$$

2.2 Unifing Item Profiles of Content-Based Filtering into the Rating Matrix

Similar with user profiles, item profiles record traces of user content features using items. Symbol $U_x \in$ U is determined by the formula (8) is the set of users i∈U used the item x∈P.

$$U_x = \{i \in U \mid r_{ix} \neq 0 \quad (i \in U, x \in P\ )\} \quad (8)$$

Symbol User (x, q) is the set of users i $\in U_x$ containing user content feature q∈T be determined by the formula (9).

$$User\,(x,q) = \{i \in U_x \mid t_{iq} \neq 0 \quad (i \in U, x \in P,\ q \in T)\} \quad (9)$$

Based on U_x and User(x, q), content-based recommender methods estimate weight v_xq reflecting importance of the user content feature q to the item x.

$$v_{xq} = \begin{cases} \frac{1}{|User(x,q)|} \sum_{i \in User(x,q)} r_{ix} & If\ |User\,(x,q)| \geq \theta \\ \frac{1}{\theta} \sum_{i \in User(x,q)} r_{ix} & If\ \ |User\,(x,q)| < \theta \end{cases} \tag{10}$$

The extended rating matrix according to item profiles is determined by the formula (11). Inthere, i = q (q∈T) acts as a complementary user to extend the rating matrix to users side.

$$r_{ix} = \begin{cases} r_{ix} & If\ i \in U\ and\ r_{ix} \neq 0 \\ v_{qx} & If\ \ q \in T\ and\ v_{qx} \neq 0\ (i = q) \end{cases} \tag{11}$$

3 The Semi-supervised Learning Model for Hybrid Filtering

As we mentioned above, recommender methods based on (7), (11) had problems with sparse data [1,9,11]. To remedy this, we propose a hybrid recommender algorithm by semi-supervised learning method. The algorithm is built based on two semi-supervised procedures: the first procedure semi-supervise the set of ratings between users and item's features, the second procedure semi-supervise the set of ratings between items and users features. The first procedure allows us to detect new items that is high suitable capability with the users. New predicted items will transfer to the process of semi-supervised according item ratings and the set of user features. The second procedure allows us to detect new users that is high suitable ability with the items. New predicted users will tranfer to the process of semi-supervised according user ratings and the set of item features. Two procedures performed simultaneously and complement each other for suitable predicted values to improve recommender results.

3.1 Semi-supervise the Set of Ratings Between Users and Item's Features

The hybrid recommender system determined by (7) allows us to deploy easily collaborative filtering methods based on users [7,10,11]. The method is conducted through 4 steps: calculating similarities between pairs of users, determining set of neighbors with the active users, predict users opinion with new items and recommend top-k items having highest predictive values for the active user [7,11]. In the first step, with each user i∈U, we build the set Si be determined by the formula (12) to observe the calculation of similarities between pairs of users.

$$S_i = \{j \in U : |P_i \cap P_j| \geq \theta_1 v\grave{a}\ \ |C_i \cap C_j| \geq \theta_2\} \tag{12}$$

$$C_i = \{s \in C : r_{is} \neq 0\} \tag{13}$$

The S_i is determined by (12) is the set of users j∈U have number of intersectant ratings with the user i at least θ_1 items and number of intersectant item

features at least θ_2. Two positive integer constants θ_1 and θ_2 are selected big enough in the training data set to S_i no longer sparse data set. Based on S_i and Pearson correlation, we semi-supervise the calculation similarities between pairs of users in collaborative filtering by the formula (14), semi-supervise the calculation similarities between pairs of users in content-based filtering by the formula (15), semi-supervise the calculation similarities between pairs of users in hybrid filtering by the formula (16).

$$a_{ij} = \begin{cases} 0 & If\ j \notin S_i \\ \frac{\sum_{x \in P_i \cap P_j} (r_{ix} - \bar{r_i})(r_{jx} - \bar{r_j})}{\sqrt{\sum_{x \in P_i \cap P_j} (r_{ix} - \bar{r_i})^2} \sqrt{\sum_{x \in P_i \cap P_j} (r_{jx} - \bar{r_j})^2}} & If\ j \in S_i \end{cases} \tag{14}$$

$$b_{ij} = \begin{cases} 0 & If\ j \notin S_i \\ \frac{\sum_{s \in C_i \cap C_j} (r_{is} - \ddot{r_i})(r_{js} - \ddot{r_j})}{\sqrt{\sum_{s \in C_i \cap C_j} (r_{is} - \ddot{r_i})^2} \sqrt{\sum_{s \in C_i \cap C_j} (r_{js} - \ddot{r_j})^2}} & If\ j \in S_i \end{cases} \tag{15}$$

$$u_{ij} = \begin{cases} \frac{\sum_{x \in H_i \cap H_j} \left(r_{ix} - \overline{\overline{r_i}}\right)\left(r_{jx} - \overline{\overline{r_j}}\right)}{\sqrt{\sum_{x \in H_i \cap H_j} \left(r_{ix} - \overline{\overline{r_i}}\right)^2} \sqrt{\sum_{x \in H_i \cap H_j} \left(r_{ix} - \overline{\overline{r_i}}\right)^2}} & If\ j \in S_i, a_{ij} \geq \alpha, b_{ij} \geq \alpha \\ 0 & other\ case \end{cases} \tag{16}$$

Inthere, H_i, $\bar{r_i}$, $\ddot{r_i}$, $\overline{\overline{r_i}}$ are determined by the formula (17), (18), (19) and (20)

$$H_i = P_i \cup C_i \tag{17}$$

$$\bar{r_i} = \frac{1}{|P_i \cap P_j|} \sum_{x \in P_i \cap P_j} r_{ix} \tag{18}$$

$$\ddot{r_i} = \frac{1}{|C_i \cap C_j|} \sum_{s \in C_i \cap C_j} r_{is} \tag{19}$$

$$\overline{\overline{r_i}} = \frac{1}{|H_i \cap H_j|} \sum_{x \in H_i \cap H_j} r_{ix} \tag{20}$$

Clearly, a_{ij} is determined on S_i by (14) more accurately than a_{ij} be determined on total the user set U in the training data set because S_i project on item columns is not sparse dataset. The b_{ij} value is determined on S_i by (15) more accurately than b_{ij} be determined on total item features C because S_i project on item feature columns is not sparse dataset. The value u_{ij} is determined by (16) is trustier than u_{ij} be determined on total user set because S_i isn't sparse dataset on total U∪C. Futhermore, two user i,j have similarity according users ratings and similarity according user profiles must exceed a certain threshold α. After determining similarities between pairs of users, we build neighbors set of the user i∈U following the formula (21). The predictive method new items x∈P that unknown by the user i be performed by the formula (22) [7,11,12].

$$K_i = \{j \in S_i : \quad u_{ij} > \alpha\} \tag{21}$$

$$r_{ix} = \overline{\overline{r_i}} + \frac{\sum_{j \in K_i} \left(r_{jx} - \overline{\overline{r_j}}\right) u_{ij}}{\sum_{j \in K_i} |u_{ij}|} \tag{22}$$

New items x∈P having the rating value r_{ix} are trusty prediction that be supplemented into the extended rating matrix according to item profiles to semi-supervise the set of ratings between items and user's features.

3.2 Semi-supervise the Set of Ratings Between Items and Users Features

Similar with Sect. 3.1, with each user x∈P, we build the set S_x be determined by the formula (23) to observe the calculation of similarities between pairs of items.

$$S_x = \{y \in P : |U_x \cap U_y| \geq \gamma 1 \; v\grave{a} \; |T_x \cap T_y| \geq \gamma 2\} \tag{23}$$

$$T_x = \{q \in T : r_{qx} \neq 0\} \tag{24}$$

Based on S_x and Pearson correlation, we semi-supervise the calculation similarities between pairs of items in collaborative filtering by the formula (25), semi-supervise the calculation similarities between pairs of items in content-based filtering by the formula (26), semi-supervise the calculation similarities between pairs of items in hybrid filtering by the formula (27).

$$a_{xy} = \begin{cases} 0 & If \; y \notin S_x \\ \dfrac{\sum_{i \in U_x \cap U_y} (r_{ix} - \bar{r_x})(r_{iy} - \bar{r_y})}{\sqrt{\sum_{i \in U_x \cap U_y} (r_{ix} - \bar{r_x})^2} \sqrt{\sum_{i \in U_x \cap U_y} (r_{iy} - \bar{r_y})^2}} & If \; y \in S_x \end{cases} \tag{25}$$

$$b_{xy} = \begin{cases} 0 & If \; y \notin S_x \\ \dfrac{\sum_{q \in T_x \cap T_y} (r_{qx} - \dddot{r_x})(r_{qy} - \dddot{r_y})}{\sqrt{\sum_{q \in T_x \cap T_y} (r_{qx} - \dddot{r_x})^2} \sqrt{\sum_{q \in T_x \cap T_y} (r_{qy} - \dddot{r_y})^2}} & If \; y \in S_x \end{cases} \tag{26}$$

$$p_{xy} = \begin{cases} \dfrac{\sum_{i \in H_x \cap H_y} (r_{ix} - \overline{\overline{r_x}})(r_{iy} - \overline{\overline{r_y}})}{\sqrt{\sum_{i \in H_x \cap H_y} (r_{ix} - \overline{\overline{r_x}})^2} \sqrt{\sum_{i \in H_x \cap H_y} (r_{iy} - \overline{\overline{r_y}})^2}} & If \; y \in S_x \; and \\ & a_{xy} \geq \alpha \; and \; b_{xy} \geq \alpha \\ 0 & Othercase \end{cases} \tag{27}$$

Inthere, H_x, $\bar{r_x}$, $\dddot{r_x}$, are determined by the formula (28), (29), (30) and (31)

$$H_x = U_x \cup T_x \tag{28}$$

$$\bar{r_x} = \frac{1}{|U_x \cap U_y|} \sum_{i \in U_x \cap U_y} r_{ix} \tag{29}$$

$$\dddot{r_x} = \frac{1}{|T_x \cap T_y|} \sum_{q \in T_x \cap T_y} r_{qx} \tag{30}$$

$$\overline{\overline{r_x}} = \frac{1}{|H_x \cap H_y|} \sum_{i \in H_x \cap H_y} r_{ix} \tag{31}$$

After determining similarities between pairs of items, we build neighbors set of the item x∈P following the formula (32). The predictive method for suitableness of the user i∈U with the item x∈P be performed by the formula (33) [8,11,12].

$$K_x = \{y \in S_x : \quad p_{xy} > \alpha\} \tag{32}$$

$$r_{ix} = \frac{\sum_{y \in K_x} p_{xy} r_{iy}}{\sum_{y \in K_x} |p_{xy}|} \tag{33}$$

3.3 The Semi-supervised Learning Algorithm for Hybrid Filtering

In this section, we propose a semi-supervised learning algorithm simultaneously to perform a transfering process for predictive results between the two methods. The algorithm is described in detail by Fig. 1.

```
Input:
  - The rating matrix R  is determined by (1)
  - The matrix of item content features C is determined by(2)
  - The matrix of user content features T is determined by(3)
  - The user i∈U need recommendation, the active user
Output : The predicted matrix
  R = R^(t) = { r_ix^(t) : i = 1, 2,.., N;  x = 1, 2,.., M }
Steps:
Begin
  Step 1( Initiative step):
    t← 0;//Init number of iteration at the beginning is 0
    R^(0) = { r_ix^(0) : i = 1, 2,.., N;  x = 1, 2,.., M }
Step 2 (Iterative step):
  Repeat
   2.1. Semi-supervise the set of ratings between users and
        item's features
      a)Determine weights of item content features at the
        iterative step t following the formula 6)
      b)Extend the rating matrix according to user profiles
        by the formula(7)
      c)Determine S_i^(t) following the formula (12)
      d)Calculate u_ij^(t) following the formula (16)
      e)Determine K_i^(t) following the formula (21)
      f)Predict r_ix^(t) following the formula (22)
```

Fig. 1. The semi-supervised learning algorithm

```
    2.2. Semi-supervise the set of ratings between items
      and user's features
      a)Determine weights of user content features at the
        iterative step t following the formula (10)
      b)Extend the rating matrix according to item profiles
        by the formula (11)
      c)Determine S_x^(t) following the formula (23)
      d)Calculate p_xy^(t) following the formula (27)
      e)Determine K_x^(t) following the formula (32)
      f)Predict r_ix^(t) following the formula (33)
2.3. Increase iterative step: t← t+1;
Until Converges.
End.
```

Fig. 1. *continued*

4 Experiment and Evaluation

4.1 Data Set

The semi-supervised learning algorithm for hybird filtering is experimented by the data set MovieLens 1M of the research group GroupLens belong to Minnesota university [13]. We choose MovieLens 1M because this dataset provide film content features and user content features fully. The data set MovieLens 1M includes 1M ratings of 6040 users, 3952 movies. Rating levels set from 1 to 5. Sparse level of the rating data is 99.1 %.

4.2 Experimental Method

At first, all experimental data set is divided into 2 parts, one part Utr is used as training data, the rest data Ute is testing data. The Utr contains 80 % ratings and Ute contains 20 % ratings. The training data is used to build model following the representative algorithm above. Each user i belongs to the testing data, the exit ratings of the active user is divided into 2 parts O_i and P_i. O_i is known, whereas P_i is ratings that need prediction from the training data and O_i [2,3,13]. To experiment the proposed methods, we evaluate the forecasting error by MAE measure. If the value MAE is small, the predictive method will give high accuracy [2,3,13].

4.3 Compararison and Evaluation

The semi-supervised learning was proposed by Sect. 3 will be experimented and compared with methods below:

- The method KNN based on user using the correlative measure Pearson (Symbol CF-UserBased) [8].
- The method KNN based on user profiles using the correlative measure Pearson (Symbol CBF-UserBased) (15).

– The method KNN based on item profiles using the correlative measure Pearson (Symbol CBF-ItemBased) (26).
– The hybrid recommender method KNN based on users and the set of item features using the correlative measure Pearson (Symbol Hybrid-UserBased) (16).
– The hybrid recommender method KNN based on items and the set of user features using the correlative measure Pearson (Symbol Hybrid-ItemBased) (27).

Through experimental process, we select $\theta = 15$, $\theta_1 = 15$, $\theta_2 = 15$ Select randomly *4000* users in the dataset MovieLens to become training dataset. Select randomly 1000 users in rest of the dataset Movielens to become testing dataset. The value *MAE* in Table 1 is averaged of 10 times of random experiments.

Table 1. MAE value of some methods

Method	Number of recommender items		
	10	20	30
CBF-UserBased	0.865	0.859	0.855
CBF-ItemBased	0.894	0.883	0.875
CF-UserBased	0.824	0.817	0.813
CF-ItemBased	0.846	0.841	0.836
Hybrid-UserBased	**0.793**	**0.792**	**0.791**
Hybrid-ItemBased	**0.798**	**0.788**	**0.782**
Semi-Learning	**0.672**	**0.629**	**0.617**

The results in Table 1 show that the method Semi-Learning give the lowest MAE in comparison with all remain methods. This can be confirmed the methods to determine similarity based on not sparse dataset with users and items are reliable completely. The method transfer predicted values between 2 processes: Semi-supervise the set of ratings between users with item's features and semi-supervise the set of ratings between items with users features to limit effectively for sparse data problem of collaborative filtering.

5 Conclusions and Developable Direction

The paper proposed the unify model between collaborative filtering recommender methods and content-based fitlering recommender methods by semi-supervised learning method. The method is performed by unifing representations of item features into collaborative filtering to unify predicted methods based on user. Then, building the unified method to represent values of user features into collaborative filtering to unify predicted methods based on items. Finally, we build the semi-supervised learning method to transfer predicted results between the two methods: give prediction according to users and give prediction according to items.

To promote the advantages and limit disadvantages of filtering methods, we propose to build two steps of supervisions: Semi-supervise the set of ratings between users with item's features and Semi-supervise the set of ratings between items and users features. The first step is performed by building non-sparse dataset with each users. The second step is performed by building non-sparse dataset with each items. Based on non-sparse dataset with each user and item, we have limited process of calculating similarities between pairs of users, neighbors set of users and items to determine reliably predicted results. On bases of two semi-supervised processes have been developed, we proposed the semi-supervised learning algorithm to transfer predicted values between semi-supervised processes. Experimental results on the real dataset about films show that the proposed methods achieve superior performance compared to baseline methods.

References

1. Su, X., Khoshgoftaar, T.M.: A survey of collaborative filtering techniques. In: Advances in Artificial Intelligence, pp. 1–20 (2009)
2. Burke, R.D.: Hybrid recommender systems: survey and experiments. User Model. User-Adapted Interact. **12**(4), 331–370 (2002)
3. Gunawardana, A., Guy, S.: A survey of accuracy evaluation metrics of recommendation tasks. J. Mach. Learn. Res. **10**, 2935–2962 (2009)
4. Burke, R., Vahedian, F., Mobasher, B.: Hybrid recommendation in heterogeneous networks. In: Dimitrova, V., Kuflik, T., Chin, D., Ricci, F., Dolog, P., Houben, G.-J. (eds.) UMAP 2014. LNCS, vol. 8538, pp. 49–60. Springer, Heidelberg (2014). doi:10.1007/978-3-319-08786-3_5
5. Pazzani, M.J.: A framework for collaborative, content-based and demographic filtering. Artif. Intell. Rev. **13**(5–6), 393–408 (1999)
6. Claypool, M., Gokhale, A., Miranda, T., Murnikov, P., Netes, D., Sartin, M.: Combining content-based and collaborative filters in an online newspaper. In: Proceedings of ACM SIGIR Workshop on Recommender Systems, vol. 60. Citeseer (1999)
7. Breese, J.S., Heckerman, D., Kadie, C.: Empirical analysis of predictive algorithms for collaborative filtering. In: Proceedings of 14th Conference on Uncertainty in Artificial Intelligence (1998)
8. Sarwar, B., Karypis, G., Konstan, J., Riedl, J.: Item-based collaborative filtering recommendation algorithms. In: Proceedings of 10th International WWW Conference (2001)
9. Phuong, N.D., Phuong, T.M.: Collaborative filtering by multi-task learning. In: RIVF 2008, pp. 227–232 (2008)
10. Lien, D.T., Anh, N.X., Phuong, N.D.: A graph model for hybrid recommender systems. In: KSE 2015, pp. 138–143 (2015)
11. Quang, T.N., Lien, D.T., Phuong, N.D.: Collaborative filtering by co-training method. In: KSE 2014, pp. 273–285 (2014)
12. Wang, J., de Vries, A.P., Reinders, M.J.T.: Unifying user-based and item-based collaborative filtering approaches by similarity fusion. In: Proceedings of the 29th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval (SIGIR 2006), pp. 501–508. ACM, New York (2006)
13. Herlocker, J.L., Konstan, J.A., Terveen, L.G., Riedl, J.T.: Evaluating collaborative filtering recommender systems. ACM Trans. Inf. Syst. **22**(1), 5–53 (2004)
14. Social Computing Research at the University of Minnesota. http://www.grouplens.org/

A Study on Fitness Representation in Genetic Programming

Thuong Pham Thi[1,3(✉)], Xuan Hoai Nguyen[2], and Tri Thanh Nguyen[3]

[1] University of Information and Communication Technology, Thainguyen University, Thai Nguyen, Vietnam
ptthuong@ictu.edu.vn
[2] Hanoi University, Hanoi, Vietnam
nxhoai@hanu.edu.vn
[3] VNU University of Engineering and Technology, Hanoi, Vietnam
ntthanh@vnu.edu.vn

Abstract. In this paper, we propose a variation on the fitness function in Genetic Programming based on Bias-Variance Genetic Programming (BVGP) [2], called BVGP*. In order to evaluate the effectiveness of this variation, we compare it with Genetic Programming [1] and Bias-Variance Genetic Programming (BVGP) [2]. The experimental results shown that the learned model by BVGP* is better than that of GP and BVGP in ability to generalize, model complexity and evaluation time.

Keywords: Genetic programming · Bias-Variance Decomposition · Regression problems

1 Introduction

Genetic Programming (GP) is one of evolutionary algorithm-based methodologies inspired by biological evolution. It uses tree-based structures and a suite of defined Genetic Algorithm-operators to generate and evolve a population of solutions to the given problem [3]. GP has produced many novel and outstanding results in various areas such as optimization, searching, sorting, quantum computing, electronic design, game playing, cyberterrorism prevention [4,6]. One of main areas of GP is Machine Learning

In Machine Learning, generalization and over-fitting are two central challenges need to be solved. Generalization error of learners directly relates to over-fitting and is referred to as the problem of over-fitting [7]. There are many researches in Machine Learning, including GP, try to improve generalization ability of learners by reducing over-fitting error as [2,8–12].

Over-fitting can be controlled by Bias-Variance trade-off [2], where bias is the error on training data set and variance is the error of difference on various data sets in the future. Over-fitting will be reduced when bias and variance are small, simultaneously. Because bias and variance are hidden in the L2-norm loss function (e.g. RMSE, MSE, ...). So, many researches in Machine Learning

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_13

have used these functions [13–16] for learning. However, the combination of bias and variance in the error function L2 sometimes causes difficulties in optimizing them simultaneously because Bias and Variance are two conflicting problems. So, in GP, Alexandros et al. proposed the method (BVGP) [2] to overcome this issue. He divided the fitness function into two components: variance and squared bias which aim at bringing variance component into the evolution process more directly. However, this method faces to the over-fit issue on limited training sample. This leads to reducing the ability to generalize of the learned model. Moreover it can make the model very sensitive to noise.

In this paper, we propose a variation on the fitness function for GP which aims at improving the limits of BVGP as shown above. It is called BVGP*. Through experiments, we demonstrate that the use of BVGP* has some advantages: (1) It can help to reduce over-fitting on the problems that GP was over-fitted; (2) the program runs faster and finds the simpler solution. So, the main contribution of this paper is the variation on the fitness function for improving the effectiveness of GP based on bias-variance decomposition of training errors.

The remainder of this paper is organized as follows: In Sect. 2, we briefly present some background knowledge and related work. A variation on the fitness function is presented in Sect. 3. Section 4 are some experimental settings and problems for testing. Next, experimental results are given in Sect. 5. Finally, Sect. 6, we summarize achieved research results and present some future works.

2 Background and Related Work

2.1 Bias-Variance Decomposition

In this section we introduce the background on the statistical concept of loss function and Bias-Variance Decomposition for regression. The material is based on the book of Trevor Hastie [17].

If we assume that $Y = f(x) + \varepsilon$, where ε is prediction error; $E(\varepsilon) = 0$ and $Var(\varepsilon) = \sigma_\varepsilon^2$, we can derive an expression for the expected prediction error of $\widehat{f}(x)$ at an input point $X = x_0$, using L2-loss function as follows:

$$\begin{aligned} Err(x_0) &= E[(Y - \widehat{f}(x))^2 | X = x_0] \\ &= \sigma_\varepsilon^2 + [E\widehat{f}(x_0) - f(x_0)]^2 + E[\widehat{f}(x_0) - E\widehat{f}(x_0)]^2 \\ &= \sigma_\varepsilon^2 + Bias^2(\widehat{f}(x_0)) + Var(\widehat{f}(x_0)) \\ &= Irreducible\ Error + Bias^2 + Variance \end{aligned} \tag{1}$$

The first term is the variance of the target around its true mean $f(x_0)$; the second term is the squared bias, the amount by which the average of our estimate differs from the true mean; the last term is the variance; the expected squared deviation of $\widehat{f}(x_0)$ around its mean. The last two terms need to be addressed for a good performance of the prediction model.

Generalization error is the prediction error over an independent test sample:

$$Err(T) = E[L(Y, \widehat{f}(x))|T] \tag{2}$$

where both X and Y are drawn randomly from their joint distribution (population). Here, the training set T is fixed, and test error refers to the error for this specific training set. A related quantity is the expected prediction error:

$$Err = E[L(Y, \widehat{f}(x))] = E[Err_T] \tag{3}$$

Such decomposition is known as the Bias Variance Decomposition.

2.2 Bias-Variance Genetic Programming (BVGP)

The Bias-Variance Genetic Programming proposed by Alexandros et al. is a new method for over-fitting issue based on Bias/Variance Error Decomposition which aims at relaxing the sensitivity of an evolved model to a particular training dataset. This method used the fitness function that is the combination of bias and variance as follows:

$$fitness = w_b Bias(D) + w_v Var(D^*) \tag{4}$$

where w_b, w_v are the coefficients for error and variance respectively; D is the training data set of size n; D^* includes B bootstrap datasets randomly drawn from D by the bootstrap re-sampling method; $Bias(D)$ is the mean error on the original dataset (bias); $Var(D^*)$ is variance of the error on the bootstrap datasets.

He separated regression error into two components: bias and variance to put variant error in the evolution process more directly.

3 The Improved Method: BVGP*

In this section, we propose a variation on the fitness function for GP which aims at overcoming the disadvantage of the BVGP. This function is based on BVGP and defined as follows:

$$fitness = w_b Bias(D^*) + w_v Var(D^*) \tag{5}$$

where bias and variance are calculated using the bootstrap re-sampling method. We consider $f(x)$ as the model trained on a dataset $D = \{(x_1, t_1), ..., (x_N, t_N)\}$ and use the bootstrap re-sampling method to randomly draw B datasets with replacement from D, each sample the same size as D. We denote D^* to include B the bootstrap sample sets: $D^* = \{D^{*b}, b : 1..B\}$. The estimated bias ($\mu$) and variance ($\sigma^2$) of stochastic fitness are computed as follows:

$$Bias(D^*) = \sum_{b=1}^{B} Bias^{*b}/B \tag{6}$$

where $Bias^{*b}$ is the bias of bootstrap sample D^{*b} is calculated using the error function RMSE:

$$Bias^{*b} = \sqrt{\frac{1}{N}\sum_{i=1}^{N}(f(x_i) - t_i)^2} \tag{7}$$

So, here we use $Bias(D^*)$ rather than the mean error on the original dataset $Bias(D)$.

$$\sigma^2 = \frac{1}{B-1}\sum_{b=1}^{B}(Bias^{*b} - Bias(D^*))^2 \tag{8}$$

As shown in [5], given a data sample, statistical inference is the process to assess how systems will behave in untested situations. It permits generalizations of conclusions beyond the sample, about an unseen population from which the sample is drawn. This process is inference from statistics to parameters, where statistics are functions on samples and parameters are functions on populations. It is noted that, $Bias(D)$ is a statistic on D while $Bias(D^*)$ is a parameter inferred from this statistic. The bootstrap re-sampling method is used to construct empirical sampling distributions for parameter estimation $Bias(D^*)$ without making any troubling assumptions about sampling models and population distributions. BVGP* learns to optimize the fitness function based on $Bias(D^*)$ while the fitness function of BVGP is based on $Bias(D)$. Therefore, we believe that the generalization ability of BVGP* is better than that of BVGP. The experimental results have confirmed this is true with most of the problems to be tested.

4 Experimental Setting

4.1 Problems

In this paper, we used benchmarks in [2] as shown in Table 1. Besides, we also used three more UCI data sets as shown in Table 2 to test the generalization ability of BVGP*. With UCI data sets, we divide an original dataset into two parts randomly: $\langle$Train sample:Test sample$\rangle = \langle 1:2 \rangle$.

4.2 GP System Setup

Evolutionary parameter values for GP systems are shown in Table 3. These typical settings are often used by GP researchers and practitioners [1].

5 Results and Discussion

In this section we present results of comparing the performance of BVGP* in comparison with GP, BVGP. We evaluate the effectiveness of BVGP* on three aspects: (1) Generalization ability; (2) Model complexity; and (3) Time complexity.

Table 1. GP benchmark regression problems

ID	Name	Definition	Training data	Testing data
BEN_1	F4	$30\frac{(x_1-1)(x_3-1)}{x_2^2(x_1-10)}$	x_1 : U[0.05, 2, 200] x_2 : U[1, 2, 200] x_3 : U[0.05, 2, 200]	x_1 : E[-0.05, 2.1, 0.15] x_2 : E[0.95, 2.05, 0.1] x_3 : E[-0.05, 2.1, 0.15]
BEN_2	F5	$6\sin(x_1)\cos(x_2)$	U[0.1, 5.9, 200]	E[−0.05, 6.05, 0.02]
BEN_3	F7	$\frac{e^{(x_1-1)^2}}{1.2+(x_2-2.5)^2}$	U[0.3, 4, 200]	E[−0.2, 4.2, 0.01]
BEN_4	F8	$x_1x_2+\sin((x_1-10)(x_2-1))$	U[−3, 3, 200]	E[−3, 3, 0.01]
BEN_5	F9	$x_1^4-x_1^3+\frac{x_2^2}{2}-x_2$	U[−3, 3, 200]	E[−3, 3, 0.01]
BEN_6	F10	$\frac{8}{2+x_1^2+x_2^2}$	U[−3, 3, 200]	E[−3, 3, 0.01]
BEN_7	F11	$\frac{x_1^3}{5}+\frac{x_2^3}{2}-x_2-x_1$	U[−3, 3, 200]	E[−3, 3, 0.01]

Table 2. UCI data sets

ID	Name	No. of attributes	#train samples	#testing samples
UCI_1	Census6	6	133	267
UCI_2	No2	7	167	333
UCI_3	SkillCraft1_Dataset	19	1114	2224

Table 3. GP systems setup

Paramaters	GP	BVGP	BVGP*
Problems	See Tables 1 and 2		
EA used in GP systems	Elitist, generational,		
	expression tree representation		
Function set	+, −, *, / (PD)		
Terminal set	Regression variables and		
	one random constant in [0.0, 1.0]		
No. of generations	151		
Population size	500		
Tournament size	4		
Tree creation	Ramped half-and-half		
	(depths of 2 to 6)		
Max. tree depth	15		
Sub tree crossover rate	0.9		
Sub tree mutation rate	0.1		
No. of Runs	100		
Fitness function	RMSE		
No. of bootstrap datasets		30	30
wb, wv		0.7, 0.3	0.7, 0.3

5.1 Generalization Error (Fittest)

In this section, we repeated one hundred runs independently for each GP system. Generalization error is the median of testing error of the best individual from all these runs. The Table 4 shows the generalization error or testing error (fittest) GP, BVGP and BVGP*, bold values indicate that the corresponding method is the best result. We see that with most of problems (BEN_1, BEN_2, BEN_3, BEN_4, BEN_5, BEN_7, UCI_1) fittest error of BVGP* is smaller than that of GP and BVGP or generalization ability of BVGP* is better than that of GP and BVGP. However, with UCI_2, generalization ability of BVGP* is much worse than that of GP and BVGP. The cause can be the learned model by BVGP is under-fit on this problem.

It is noted that both the GP and BVGP use the bias on the original training dataset ($Bias(D)$) as the optimal goal, this lead to over-fitting when the size of the training sample is limited or there is noise in the train data or the sampling process is bad. BVGP* rather than using $Bias(D)$, it uses the mean of the empirical bootstrap error distribution ($Bias(D^*)$) as one of the optimal goals. So, it can avoid the sampling bias issues that lead to over-fitting solution as showed above. This explains why the results by BVGP* are better than those of GP and BVGP in most of problems.

5.2 Model Complexity and Evaluation Time

In this section, we repeated one hundred runs independently for each GP system. Generalization error is the median of testing error of the best individual from these runs. The Table 5 shows the evaluation time and model complexity of the best individual by GP, BVGP, BVGP*. Bold values indicate that the corresponding method is the best.

Table 4. Summary of fittest error (median). Statistics based on 100 independent runs. Bold values indicate that the method is the best.

Problem	GP	BVGP	BVGP*
BEN_1	1.404	1.6015	**1.347**
BEN_2	387.5605	425.2415	**262.844**
BEN_3	10.054	9.0615	**3.7185**
BEN_4	134.748	87.4175	**0.7**
BEN_5	486.8505	458.1465	**450.7025**
BEN_6	2.788	**0.8295**	0.85
BEN_7	623.178	854.623	**580.851**
UCI_1	0.1955	0.1952	**0.195**
UCI_2	1.5175	**1.248**	580.851
UCI_3	15.002	**15**	22

Table 5. Evaluation time (median), model complexity (median) is the average number of nodes on the best individual. Statistics based on 100 independent runs. Bold values indicate the method is the best.

Problem	Evaluation time			Model complexity		
	GP	BVGP	BVGP*	GP	BVGP	BVGP*
BEN_1	372.50	269.50	**262.50**	204.89	**197.44**	198.49
BEN_2	112.50	68.00	**65.50**	293.57	292.53	**252.92**
BEN_3	70.00	34.00	**3.50**	155.76	151.95	**4.56**
BEN_4	53.00	14.00	**2.00**	218.64	192.74	**4.91**
BEN_5	59.50	**17.00**	18.50	244.66	240.60	**239.26**
BEN_6	44.00	**2.00**	4.00	21.60	**3.21**	91.15
BEN_7	52.50	15.02	**15.00**	**207.94**	209.42	208.10
UCI_1	64.00	4.50	**4.00**	13.71	3.32	**1.40**
UCI_2	162.50	129.00	**17.00**	251.90	267.11	**208.10**
UCI_3	140.00	47.21	**47.00**	37.89	17.89	**17.39**

Here, the evaluation time is measured in milliseconds. It is effected mainly by model complexity. Similar to fittest error, in all problems (10/10 problems, see bold lines), BVGP* is faster than GP since it leaned the smaller model (see corresponding lines at the column Evaluation time). Comparing to BVGP, BVGP* also learned the model with smaller complexity with most of problems (7/10 problems, see bold lines), so it is faster than BVGP or evaluation time is smaller. It is noted that, on BEN_6, the model complexity of BVGP* is larger

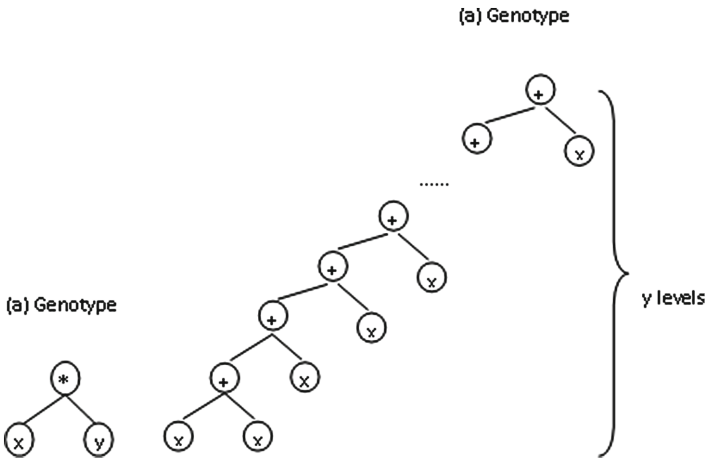


Fig. 1. The evaluation time of genotype (a) is similar to that of genotype (b), but their different model complexities are different.

while the its evaluation time is smaller than that of other methods. It can be caused by genotype of the learned model by BVGP* contains various operators that effect to evaluation time, i.e., considering two genotypes as shown in Fig. 1, although the model complexities of them are different, the evaluation time of them are similar.

6 Conclusion and Future Work

In this paper, we proposed the variation on the fitness function (BVGP*). It is based on the bias-variance decomposition and the method BVGP. Analyses of empirical results show that this approach has some advantages: (1) BVGP* can help reduce over-fitting on the problems that GP and BVGP were over-fitted; (2) It runs faster with simpler solution.

There are several future research directions arisen from this paper. First we need a more natural fitness representation way in bringing two components directly into the process of evolution. Second, we need a new selection mechanism corresponding to this fitness representation.

References

1. Koza, J.R.: Genetic Programming: on the Programming of Computers by Means of Natural Selection. MIT press, Cambridge (1992)
2. Agapitos, A., Brabazon, A., O'Neill, M.: Controlling overfitting in symbolic regression based on a bias/variance error decomposition. In: Coello, C.A.C., Cutello, V., Deb, K., Forrest, S., Nicosia, G., Pavone, M. (eds.) PPSN 2012. LNCS, vol. 7491, pp. 438–447. Springer, Heidelberg (2012). doi:10.1007/978-3-642-32937-1_44
3. Cramer, N.L.: A representation for the adaptive generation of simple sequential programs. In: Proceedings of the First International Conference on Genetic Algorithms, pp. 183–187 (1985)
4. Nordin, P.: Genetic programming iii-darwinian invention and problem solving. Evol. Comput. **7**, 451–453 (1999)
5. Cohen, P.R.: Empirical Methods for Artificial Intelligence, vol. 139. MIT press, Cambridge (1995)
6. Hansen, J.V., Lowry, P.B., Meservy, R.D., McDonald, D.M.: Genetic programming for prevention of cyberterrorism through dynamic and evolving intrusion detection. Decis. Support Syst. **43**, 1362–1374 (2007)
7. Bishop, C.M.: Pattern Recognition and Machine Learning. Springer, New York (2006)
8. Fitzgerald, J., Azad, R., Ryan, C.: A bootstrapping approach to reduce over-fitting in genetic programming. In: Proceedings of the 15th Annual Conference Companion on Genetic and Evolutionary Computation, pp. 1113–1120. ACM (2013)
9. Gonçalves, I., Silva, S., Melo, J.B., Carreiras, J.M.B.: Random sampling technique for overfitting control in genetic programming. In: Moraglio, A., Silva, S., Krawiec, K., Machado, P., Cotta, C. (eds.) EuroGP 2012. LNCS, vol. 7244, pp. 218–229. Springer, Heidelberg (2012). doi:10.1007/978-3-642-29139-5_19

10. Gonçalves, I., Silva, S.: Balancing learning and overfitting in genetic programming with interleaved sampling of training data. In: Krawiec, K., Moraglio, A., Hu, T., Etaner-Uyar, A.Ş., Hu, B. (eds.) EuroGP 2013. LNCS, vol. 7831, pp. 73–84. Springer, Heidelberg (2013). doi:10.1007/978-3-642-37207-0_7
11. Nguyen, T.H., Nguyen, X.H., McKay, B., Nguyen, Q.U.: Where should we stop? An investigation on early stopping for GP learning. In: Bui, L.T., Ong, Y.S., Hoai, N.X., Ishibuchi, H., Suganthan, P.N. (eds.) SEAL 2012. LNCS, vol. 7673, pp. 391–399. Springer, Heidelberg (2012). doi:10.1007/978-3-642-34859-4_39
12. Uy, N.Q., Hien, N.T., Hoai, N.X., O'Neill, M.: Improving the generalisation ability of genetic programming with semantic similarity based crossover. In: Esparcia-Alcázar, A.I., Ekárt, A., Silva, S., Dignum, S., Uyar, A.Ş. (eds.) EuroGP 2010. LNCS, vol. 6021, pp. 184–195. Springer, Heidelberg (2010). doi:10.1007/978-3-642-12148-7_16
13. Muttil, N., Chau, K.-W.: Neural network and genetic programming for modelling coastal algal blooms. Int. J. Environ. Pollut. **28**, 223–238 (2006). Inderscience Publishers
14. Archetti, F., Lanzeni, S., Messina, E., Vanneschi, L.: Genetic programming for human oral bioavailability of drugs. In: Proceedings of the 8th Annual Conference on Genetic and Evolutionary Computation, pp. 255–262. ACM (2006)
15. Juang, C.-F.: A hybrid of genetic algorithm and particle swarm optimization for recurrent network design. IEEE Trans. Syst. Man Cybern. Part B Cybern. **34**, 997–1006 (2004)
16. Whigham, P.A., Crapper, P.F.: Time series modelling using genetic programming: an application to rainfall-runoff models. In: Advances in Genetic Programming, vol. 3, pp. 89–104. MIT Press, Cambridge (1999)
17. Hastie, T., Tibshirani, R., Friedman, J., Franklin, J.: The Elements of Statistical Learning: Data Mining, Inference and Prediction. Springer, New York (2005). The Mathematical Intelligencer, 27, 83–85. Springer

Adaptive Robust Ability of High Order Sliding Mode Control for a 3-D Overhead Crane System

Dao Phuong Nam(✉), Nguyen Doan Phuoc, and Nguyen Thi Viet Huong

School of Electrical Engineering,
Hanoi University of Science and Technology, Hanoi, Vietnam
nam.daophuong@hust.edu.vn

Abstract. Traditionally, 3-D overhead crane systems are widely used in industry and automatic operation would reduce the risk. It is difficult to precisely position the payload in overhead crane due to the lack of actuators in this system. This paper develops an adaptive robust ability of high – order sliding mode controller (HOSMC). The finite time stability of the closed-loop system is proved without traditional Lyapunov theory. The results based on suitable second-order sliding surface and super – twisting controller. Simulation studies are performed to demonstrate the validity of the proposed control scheme.

1 Introduction

Over the past three decades, extensive research has been performed toward high performance load transportation of the overhead cranes operation. The crane is naturally an under-actuated mechanical system, in which the number of independent actuators (inputs) is less than the degree of freedom (outputs) to be controlled. So that in order to meet high performance control requirements is difficult task, naming the suitable motion speed with accurate load positioning and maintaining small low swings [1]. A number of control approaches have been suggested and mainly based on nonlinear dynamic models developed for 3D overhead cranes in adaptive – robust control [2–6]. In [4], a complete nonlinear dynamic model of an overhead crane has been proposed, and the backstepping technique that achieves 3D position control and anti-sway control simultaneously is derived in a unified control scheme under parameter variations. In [6], a nonlinear controller for payload positioning and swing suppression of overhead crane systems has been proposed and the stability analysis was performed under much less strict assumptions. Although the numerical complexity of model predictive control (MPC), a new control approach for anti-swing tracking control of a 3D overhead crane based on MPC and computed torque control is pointed out in [3], including external disturbances on the actuators driving the crane. In [13], a novel adaptive control scheme with the use of the tuning function, including both the cart motion and the swing angle dynamics is designed to ensure the stability of the closed-loop system. Some researchers implemented image sensing to measure the swing angle of an overhead crane. This work proposes the use of

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_14

visual tracking technology create the rapid and smooth load motion [8]. Sliding mode control (SMC) is capable of controlling nonlinear system with parameter uncertainties (such as weight, moment of inertial of payload, . . .) and external disturbances (such as winds), and fuzzy logic control (FLC) is independent of system model. Park et al. [7] proposed an adaptive fuzzy sliding-mode control (AFSMC) law for the trajectory tracking of 2-D overhead crane systems, subject not only to system uncertainties but also to actuator nonlinearity of the dead-zone type. In [11], an adaptive sliding mode fuzzy control algorithm based on combining SMC's robustness and FLC's independence of system model is derived. However, it is difficult to apply SMC for mechanical systems because of the sensitivity of these systems to chattering. Higher-order sliding mode control (HOSMC) can overcome this phenomenon by confining the switching control to the higher derivatives of the control variable. Bartolini et al. [9] proposed a control scheme guarantees a fast and precise load transfer and the swing suppression, based on second-order sliding surface. In [12], 2nd-order and 3rd-order sliding mode differentiators are used and actuator fault diagnosis schemes are derived to achieve fault detection and isolation. In [15], super-twisting algorithm is one of the development of high-order sliding mode control with attractive properties: finite convergence time, disturbance rejection. In [13], an separation principle output feedback controller for a class of MIMO nonlinear systems has been implemented based on HOSMC. Pisano et al. [16] proposed an anti-swing control law, which is based on the super-twisting approach for the 3-dimensional overhead crane. Le Anh Tuan et al. [2] developed an adaptive version of sliding mode controller for 3D overhead cranes. However, the Barbalat's Lemma - based proof presented in that paper was incorrect because of the lack of uniformly continuous property of. Moreover, the condition of matrices and α in sliding surface's expression was not pointed out clearly. In this paper, we propose a solution for the above problems based on uncertain model and the finite convergence time of super - twisting controller is estimated (Fig. 1).

Heading level	
l	The adjustable rope length (m)
x, z	The displacements of trolley and bridge, respectively (m)
ϕ, θ	The cargo swing angles (rad)
$m_c, m_b, m_t.m_l$	Equivalent masses of the cargo, trolley, bridge and hoist, respectively (kg)
g	Gravity acceleration (m/s^2)
u_b, u_t, u_l	The external forces driving bridge, trolley and cargo hoist, respectively (N)

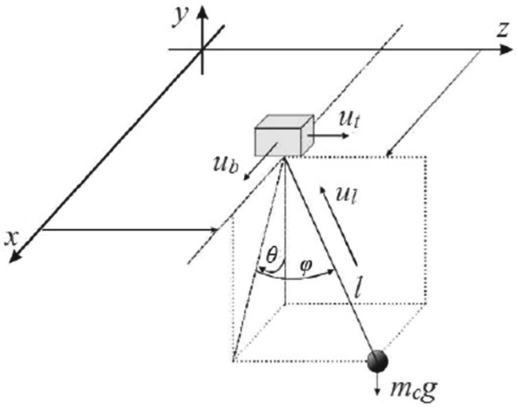


Fig. 1. 3D crane physical model [16].

2 Main Contents

2.1 3D Crane Model

The motion equations of 3D overhead crane system are created by using Lagrange's equation and can be represented in the following matrix form [2]:

$$M(q,\Theta)\ddot{q} + C(q,\dot{q},\Theta) + g(q,\Theta) = u' \tag{1}$$

where $M(q,\Theta)$ is the inertia matrix, $C(q,\dot{q},\Theta)$ is the Centripetal-Coriolis term, $g(q,\Theta)$ is vector due to gravity, and $u' \in \mathbb{R}^{n\times n}$ is the input vector. The details of the above dynamics are given in the following expressions: $q = \begin{bmatrix} z\ x\ l\ \varphi\ \theta \end{bmatrix}^T$ $u' = \begin{bmatrix} u_b\ u_t\ u_l\ 0\ 0 \end{bmatrix}^T$.

$$M = \begin{bmatrix} m_{11} & 0 & m_{13} & m_{14} & m_{15} \\ 0 & m_{22} & m_{23} & 0 & m_{25} \\ m_{31} & m_{32} & m_{33} & 0 & 0 \\ m_{41} & 0 & 0 & m_{44} & 0 \\ m_{51} & m_{52} & 0 & 0 & m_{55} \end{bmatrix} \begin{cases} m_{11} = m_t + m_b + m_c, \\ m_{13} = m_{31} = m_c \sin\varphi\cos\theta, \\ m_{14} = m_{41} = m_c l\cos\varphi\cos\theta, \\ m_{15} = m_{51} = -m_c l\sin\varphi\sin\theta, \\ m_{22} = m_t + m_c, \\ m_{23} = m_{32} = m_c \sin\theta, \\ m_{25} = m_{52} = m_c l\cos\theta, \\ m_{33} = m_l + m_c, \\ m_{44} = m_c l^2\cos^2\theta, \\ m_{55} = m_c l^2 \end{cases} \tag{2}$$

$$C = \begin{bmatrix} 0 & 0 & c_{13} & c_{14} & c_{15} \\ 0 & 0 & c_{23} & 0 & c_{25} \\ 0 & 0 & 0 & c_{34} & c_{35} \\ 0 & 0 & c_{43} & c_{44} & c_{45} \\ 0 & 0 & c_{53} & c_{54} & c_{55} \end{bmatrix} \begin{cases} c_{13} = m_c \cos\varphi \cos\theta\dot{\varphi} - m_c \sin\varphi \sin\theta\dot{\theta}, \\ c_{14} = m_c \cos\varphi \cos\theta\dot{l} - m_c l \cos\varphi \sin\theta\dot{\theta} - m_c l \sin\varphi \cos\theta\dot{\varphi}, \\ c_{15} = -m_c l \cos\varphi \sin\theta\dot{\varphi} - m_c \sin\varphi \sin\theta\dot{l} - m_c l \sin\varphi \cos\theta\dot{\theta}, \\ c_{23} = m_c \cos\theta\dot{\theta}, \\ c_{25} = m_c \cos\theta\dot{l} - m_c l \sin\theta\dot{\theta}, \\ c_{34} = -m_c l \cos^2\theta\dot{\varphi}, \\ c_{35} = -m_c l\dot{\theta}, \\ c_{43} = m_c l \cos^2\theta\dot{\varphi}, \\ c_{44} = m_c l \cos^2\theta\dot{l}, \\ c_{45} = -m_c l^2 \cos\theta \sin\theta\dot{\varphi}, \\ c_{53} = m_c l\dot{\theta}, \\ c_{54} = m_c l^2 \cos\theta \sin\theta\dot{\varphi}, \\ c_{55} = m_c l\dot{l}. \end{cases} \tag{3}$$

$$g = \begin{bmatrix} 0 & 0 & g_3 & g_4 & g_5 \end{bmatrix}^T \begin{cases} g_3 = -m_c g \cos\varphi \cos\theta, \\ g_4 = m_c g l \sin\varphi \cos\theta, \\ g_5 = m_c g l \cos\varphi \cos\theta. \end{cases} \tag{4}$$

In this paper, the dynamic Eq. (5) is presented to account for the parameter uncertainties and external disturbances:

$$M(q,d)\ddot{q} + C(q,\dot{q},d) + g(q,d) = \begin{bmatrix} u + \xi(q,\dot{q},\ddot{q},t) \\ 0 \end{bmatrix} \tag{5}$$

Where Θ is the vector of unknown constant parameters, d is the estimates for Θ and $\xi(q,\dot{q},\ddot{q}) = n(t) + \Lambda$: $n(t)$ is the vector of external disturbances; $\Lambda = (M(q,\Theta)\ddot{q} + C(q,\dot{q},\Theta)\dot{q} + g(q,\Theta)) - (M(q,d)\ddot{q} + C(q,\dot{q},d)\dot{q} + q(q,d))$ is estimation error vector of this model.

2.2 High-Order Sliding Mode Controller Design

The dynamic equation of an overhead crane (5) can be rewritten in following form by separating the model into actuated and un-actuated part 1:

$$\overline{M}(q)\ddot{q_1} + \overline{C}_1(q,\dot{q})\dot{q_1} + \overline{C}_2(q,\dot{q})\dot{q_2} + \overline{g}(q) = u + \xi \tag{6}$$

where: $q_1 = q_a = \begin{bmatrix} z & x & l \end{bmatrix}^T$ (actuated states) $q_2 = q_u = \begin{bmatrix} \varphi & \theta \end{bmatrix}^T$(unactuated states) $u = \begin{bmatrix} u_b & u_t & u_l \end{bmatrix}^T$.

The tracking control structure is designed through hign-order sliding mode control based on second-order sliding surface as follows:

$$\begin{cases} s = \dot{q_1} + \lambda + \tilde{q_1} + \alpha q_2 = 0 \\ \dot{s} = \ddot{q_1} + \lambda\dot{q_1} + \alpha\dot{q_2} = 0 \end{cases} \tag{7}$$

where $\tilde{q_1} = q_1 - q_{1d}$; $\tilde{q_2} = q_2 - q_{2d}$;
$\lambda = \begin{bmatrix} \lambda_1 & 0 & 0 \\ 0 & \lambda_2 & 0 \\ 0 & 0 & \lambda_3 \end{bmatrix}$; $\alpha = \begin{bmatrix} \alpha_1 & 0 \\ 0 & \alpha_2 \\ 0 & 0 \end{bmatrix}$.

Let us now find the control input $u_N = u_{eq} + \underline{\Lambda}$ to stabilize sliding surface in finite time. The equivalent input u_{eq} is obtained from $\dot{s} + \lambda s = 0$:

$$u_{eq} = \overline{C}_1(q,\dot{q})\dot{q}_1 + \overline{C}_2(q,\dot{q})\dot{q}_2 + g(q) - \overline{M}(q)(2\lambda\dot{q}_1 + \lambda^2\tilde{q}_1 + \alpha\dot{q}_2 + \lambda\alpha q_2) \quad (8)$$

So that, the control input $u_N = u_{eq} + \underline{\Lambda}$ obtain the following result:

$$\dot{s} + \lambda s = \overline{M}(q)^{-1}\underline{\Lambda} \quad (9)$$

If $\underline{\Lambda} = -K.sgn(s)$ with $K = diag(k_1, k_2, k_3)$ and $k_i > \delta(\forall i = 1,2,3)$ are sufficiantly large constants with $\delta = \sup\xi(q,\dot{q},\ddot{q},t)$ then it is found that s converge to 0 faster than the root of following equation:

$$\dot{s} + \overline{M}(q)^{-1}Ksgn(s) = 0 \quad (10)$$

Therefore s converge to 0 in finite time and combine with (9), we obtain $s, \dot{s}$ converge to 0 in finite time.

Remark 1: The stability of high order sliding surface is established in [2]. *However, in order to achieve a tracking performance based on sliding mode control, we need to guarantee $s, \dot{s}$ converge to 0 in finite time even with parameter uncertainties and external disturbances.*

The indentification of two metrices λ, α to make $\tilde{q}_1, \tilde{q}_2$ converge to 0 are implemented by considering the stability of following nonliner system:

$$\dot{x} = \begin{bmatrix} -\lambda x_1 - \alpha x_2 \\ x_3 \\ h(x) \end{bmatrix} = f(x) \quad (11)$$

$$x = \begin{bmatrix} x_1 & x_2 & x_3 \end{bmatrix}^T = \begin{bmatrix} \tilde{q}_1 & q_2 & \dot{q}_3 \end{bmatrix}^T$$

$$h_1(x) = -\frac{\cos\varphi}{l\cos\theta}\lambda_1^2\tilde{z} + \frac{\dot{\varphi}}{l}\lambda_3\tilde{l} - \frac{\cos\varphi}{l\cos\theta}\lambda_1\alpha_1\varphi + \left(\frac{\cos\varphi}{l\cos\theta}\alpha - \frac{\dot{l}}{l} + \tan\theta\dot{\theta}\right)\dot{\varphi}$$
$$+ \tan\theta\dot{\varphi}\dot{\theta} - \frac{g\sin\varphi}{l\cos\theta}$$
$$h_2(x) = \frac{\sin\varphi\sin\theta}{l}\lambda_1^2\tilde{z} - \frac{\cos\theta}{l}\lambda_2^2\tilde{x} + \frac{\dot{\theta}}{l}\lambda_3\tilde{\lambda} - \frac{\cos\theta}{l}\lambda_2\alpha_2\theta - \frac{\sin\varphi\sin\theta}{l}\alpha_1\dot{\varphi}$$
$$- \cos\theta\sin\theta\dot{\varphi}^2 + \frac{\cos\theta\alpha_2 - \dot{l}}{l}\dot{\theta} - \frac{q\cos\varphi\sin\theta}{l}$$

The linearization of (11) around the equilibrium 0 is given by (12):

$$\dot{x} = Ax \quad (12)$$

where

$$A = \left.\frac{\partial f}{\partial x}\right|_{x=0} = \begin{bmatrix} -\lambda & -a & 0_{3\times2} \\ 0_{2\times3} & 0_{2\times2} & I_{2\times2} \\ A_{31} & A_{32} & A_{33} \end{bmatrix} \quad (13)$$

In [14], the system (10) is local stable if this equivalent nonlinear system is also stable. After some manipulations, we get:

$$|sI_7 - A| = (s+\lambda_3)\left(s^3 + \left(\lambda_1 - \frac{\alpha_1}{l_d}\right)s^2 + \frac{g}{l_d}s + \frac{\lambda_1 g}{l_d}\right)\left(s^3 + \left(\lambda_2 - \frac{\alpha_2}{l_d}\right)s^2 + \frac{g}{l_d}s + \frac{\lambda_2 g}{l_d}\right)$$

Let:

$$P(s) = \left(s^3 + \left(\lambda_1 - \frac{\alpha_1}{l_d}\right)s^2 + \frac{g}{l_d}s + \frac{\lambda_1 g}{l_d}\right)$$
$$Q(s) = \left(s^3 + \left(\lambda_2 - \frac{\alpha_2}{l_d}\right)s^2 + \frac{g}{l_d}s + \frac{\lambda_2 g}{l_d}\right)$$

$P(s)$ and $Q(s)$ are Hurwitz polynomials if the following conditions are satisfied using Routh–Hurwitz criteria:

$$\begin{aligned} &\lambda_1 - \frac{\alpha_1}{l_d} > 0; \frac{g}{l_d} > 0; \frac{\lambda_1 g}{l_d} > 0; \left(\lambda_1 - \frac{\alpha_1}{l_d}\right)\frac{g}{l_d} > \frac{\lambda_1 g}{l_d} \\ &\lambda_2 - \frac{\alpha_2}{l_d} > 0; \frac{g}{l_d} > 0; \frac{\lambda_2 g}{l_d} > 0; \left(\lambda_2 - \frac{\alpha_2}{l_d}\right)\frac{g}{l_d} > \frac{\lambda_2 g}{l_d} \end{aligned} \tag{14}$$

For a general result, the proposed control law parameters satisfying (15) will guarantee the stability of the 3D Crane system.

$$\lambda_1 > 0, \lambda_2 > 0, \lambda_3 > 0, \alpha_1 < 0, \alpha_2 < 0 \tag{15}$$

Remark 2: Note that the control input $u_N = u_{eq} + \underline{\Delta}$ (8, 9) is not enough for complete control signal. It is necessary to choose two suitable matrices λ, α (14, 15) and it has not been pointed out in [2].

2.3 Super-Twisting Controller Design

The motion of the 3-D overhead crane (1) can be simplified by assuming that the load swing angles ϕ, θ are small enough, it can be simplified as:

$$\ddot{x} = -\frac{D_x}{M_x}\dot{x} + \frac{1}{M_x}u_b; \ddot{\theta} = \frac{D_x}{M_x}\frac{\dot{x}}{l} - \frac{1}{M_x}\frac{u_b}{l} \tag{16}$$

$$\ddot{y} = -\frac{D_y}{M_y}\dot{y} + \frac{1}{M_y}u_t; \ddot{\phi} = \frac{D_y}{M_y}\frac{\dot{y}}{l} - \frac{1}{M_y}\frac{u_t}{l}; \ddot{l} = g - \frac{D_l}{m}\dot{l} + \frac{1}{m}u_l \tag{17}$$

In [16], the tracking control structure is designed through the following three dimensional sliding variable:

$$s = \begin{bmatrix} s_x \\ s_y \\ s_l \end{bmatrix} = \begin{bmatrix} \dot{x} - \dot{x}^r + c_x(x - x^r) - k_x\theta_x \\ \dot{y} - \dot{y}^r + c_y(y - y^r) - k_y\theta_y \\ \dot{l} - \dot{l}^r + c_l(l - l^r) \end{bmatrix} \tag{18}$$

where c_x, c_y, c_l, k_x, k_y positive constants. So that, the sliding dynamics along the model (16, 17) is:

$$\frac{d}{dt}s = \begin{bmatrix} a_1(x,\dot{x},t) + b_1(x,\dot{x},t)u_b \\ a_2(y,\dot{y},t) + b_2(y,\dot{y},t)u_t \\ a_3(l,\dot{l},t) + b_3(l,\dot{l},t)u_l \end{bmatrix} \tag{19}$$

The control inputs u_b, u_t, u_l can be obtained as follows [16]:

$$\begin{aligned} &u_{b,t,l} = -\lambda_i\sqrt{|s_i|}sign(s_i) + \omega_i \\ &\dot{\omega}_i = -\alpha_i sign(s_i) \\ &(i \in x, y, l) \end{aligned} \tag{20}$$

where λ_i and α_i are sufficiently large constants.

From (19) and combine with the results (Levant 2005), the convergence time is estimated as follows:

$$T \leq \sum \frac{|\dot{x}_i|}{K_m\alpha - C}$$
$$\left(x_i = x, y, l; |\dot{a}| + U_m\left|\dot{b}\right| \leq C; 0 \leq K_m \leq b(t,x) \leq K_M; \left|\frac{a}{b}\right| < qU_M; 0 < q < 1\right) \tag{21}$$

Remark 3: The estimation of convergence time has not been pointed out in [16] *and* (21) *is the additional result in finite time stability.*

2.4 Verification by Simulation

In this section, simulations via MATLAB/Simulink are performed to verify the validity of controller (8, 9) and (20) with 2 suitable matrices λ, α satisfying (14, 15):

$$\lambda_1 = \lambda_2 = 0.75; \lambda_3 = 1; \alpha_1 = \alpha_2 = -4$$

The matrix K is chosen big enough for a fast convergence to sliding surface :

$$K = \begin{bmatrix} 0.5 & 0 & 0 \\ 0 & 0.5 & 0 \\ 0 & 0 & 1 \end{bmatrix}$$

The simulation results are shown in Figs. 2, 3 and 4. All components of sliding surface variable are controlled to converge to 0 in a finite time (Figs. 2 and 4). Hence, it is not only rapid to track with short settling time and no overshoot of the actuated states (Figs. 2, 3 and 4), but also two unactuated cargo-swing angles are also maintained small (Figs. 2, 3 and 4).

Besides, there is a small difference between parameters used for controller and plant to prove the robustness of the proposed sliding mode control law. These parameters are listed in the following table:

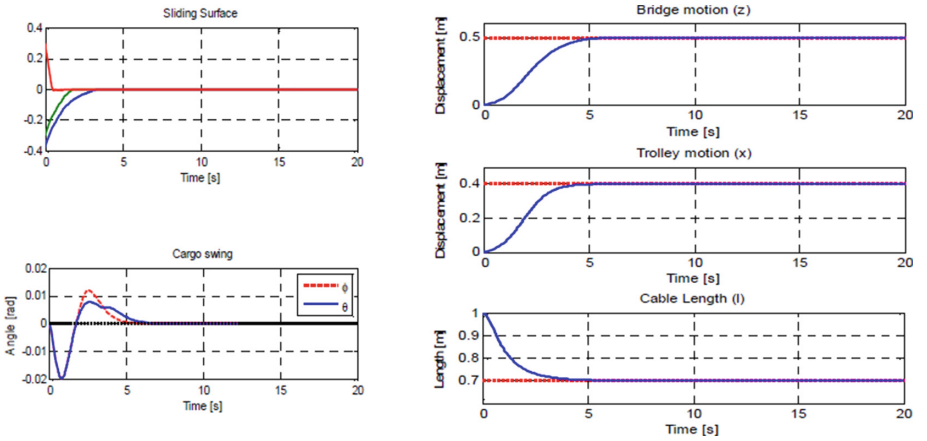


Fig. 2. Sliding surface s, Actuated states and Un-actuated states.

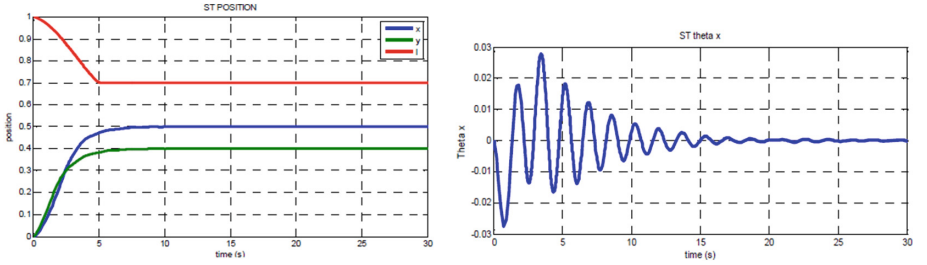


Fig. 3. Actuated states and Un-actuated states (Super – Twisting Controller).

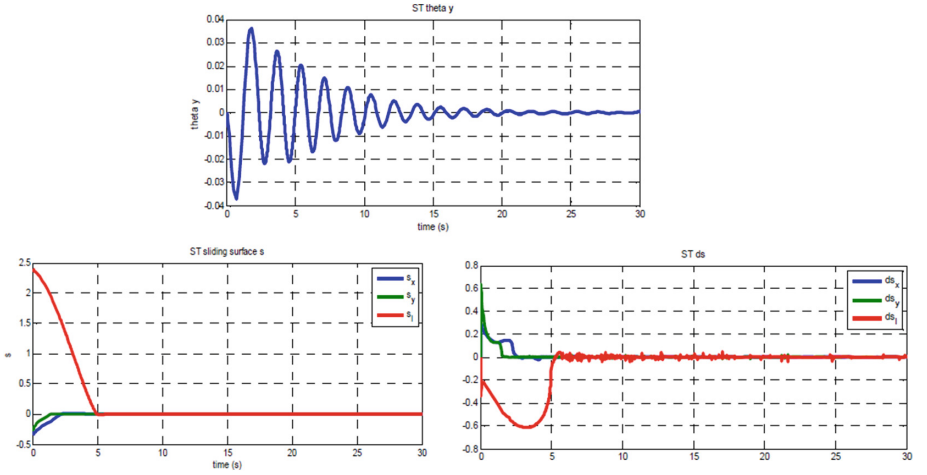


Fig. 4. Un-actuated states, sliding surface **s**, time derivative of the sliding surface **s** (Super – Twisting Controller).

Parameters	Unit	Controller	Plant
m_c	Kg	0.85	0.85
m_b	Kg	7	6.9
m_t	Kg	5	4.8
m_l	Kg	2	2.1
b_t	Kg/s	20	21
b_b	Kg/s	30	31.5
b_r	Kg/s	50	46.5

3 Conclusion

In this paper, tracking ability and finite time stability is analyzed, refer to differential equation without traditional Lyapunov theory and the use of two suitable matrices in the second-order sliding surface is proposed. It is theoretically proved that asymptotic tracking of the payload position and regulation of the swing angle can be achieved robustly despite the parameter uncertainties, external disturbance. Good results in offline simulation showed the effectiveness of the proposed theoretical development in this work.

References

1. Moustafa, K.A.F., Ebeid, A.M.: Nonlinear modeling and control of overhead crane load sway. ASME Trans. Dynamic Syst. Measur. Control **110**, 266–271 (1988)
2. Tuan, L.A., Kim, J.-J., Lee, S.-G., Lim, T.-G., Nho, L.C.: Second-order sliding mode control of a 3D overhead crane with uncertain system parameters. Int. J. Precis. Eng. Manuf. **15**(5), 811–819 (2014)
3. Khatamianfar, A., Savkin, A.V.: A new tracking control approach for 3D overhead crane systems using model predictive control. In: European Control Conference (ECC) (2014)
4. Tsai, C.-C., Lang, W.H., Chuang, K.-H.: Backstepping aggregated sliding-mode motion control for automatic 3D overhead cranes. In: IEEE/ASME International Conference on Advanced Intelligent Mechatronics (2012)
5. Hua, Y.J., Shine, Y.K.: Adaptive coupling control for overhead crane systems. Mechatronics **17**, 143–152 (2007)
6. Chwa, D.: Nonlinear tracking control of 3-D overhead cranes against the initial swing angle and the variation of payload weight. IEEE Trans. Control Syst. Technol. **17**(4), 876–883 (2009)
7. Park, M.-S., Chwa, D., Hong, S.-K.: Antisway tracking control of overhead cranes with system uncertainty and actuator nonlinearity using an adaptive fuzzy sliding-mode control. IEEE Trans. Ind. Electron. **55**(11), 3972–3984 (2008)
8. Lee, L.-H., Huang, C.-H., Sung-Chih, K., Yang, Z.-H., Chang, C.-Y.: Efficient visual feedback method to control a three-dimensional overhead crane. IEEE Trans. Ind. Electron. **61**(8), 4073–4083 (2011)

9. Bartolini, G., Pisano, A., Usai, E.: Second-order sliding-mode control of container cranes. Automatica **38**, 1783–1790 (2002)
10. Park, H., Chwa, D., Hong, K.-S.: A feedback linearization control of container cranes: varying rope length. Int. J. Control Autom. Syst. **5**(4), 379–387 (2007)
11. Liu, D., Yi, J., Zhao, D., Wang, W.: Adaptive sliding mode fuzzy control for a two-dimensional overhead crane. Mechatronics **15**, 505–522 (2005)
12. Lee, H.-H.: Motion planning for three-dimensional overhead cranes with high-speed load hoisting. Int. J. Control **78**(12), 875–886 (2005)
13. Yang, J.H., Shen, S.H.: Novel approach for adaptive tracking control of a 3-D overhead crane system. J. Intell. Robot. Syst. **62**(1), 59–80 (2011)
14. Phuoc, N.D.: Phan tich va dieu khien he phi tuyen, NXB Bach Khoa (2012)
15. Utkin, V.: On convergence time and disturbance rejection of super-twisting control. IEEE Trans. Autom. Control, **58**(8) (2013)
16. Pisano, A., Scodina, S., Usai, E.: Load swing suppression in the 3-dimensional overhead crane via second-order sliding-modes on convergence time and disturbance rejection of super-twisting control. In: 11th International Workshop on Variable Structure Systems Mexico City, Mexico, 26–28 June 2010

An Evaluation of Hand Pyramid Structure for Hand Representation Based on Kernels

Van-Toi Nguyen[1(✉)], Thi-Lan Le[2], and Thanh-Hai Tran[2]

[1] School of Information and Communication Technology, Thai Nguyen University, Thai Nguyen, Vietnam
nvtoi@ictu.edu.vn

[2] International Research Institute MICA, HUST-CNRS/UMI-2954-GRENOBLE INP, Hanoi University of Science and Technology, Hanoi, Vietnam
{thi-lan.le,thanh-hai.tran}@mica.edu.vn
http://www.ictu.edu.vn, http://www.mica.edu.vn

Abstract. Hand posture recognition is an active research topic in computer vision and robotics with many applications ranging from automatic sign language recognition to human-system interaction. Recently, we have proposed a new descriptor for hand representation based on the kernel method (KDES) [1]. Our new descriptor inherits the main idea of KDES but we proposed three improvements to make it more robust. One of the improvements was that we introduced a new hand pyramid structure [14]. Intuitively, hand pyramid is more suitable to hand structure than conventional pyramid. In our previous work, we have demonstrated that the combination of improvements to KDES gives more accurate hand posture classification than using original KDES. However, it still lacks discussions and experimental evidences of the contribution of hand pyramid for hand representation. In this paper, we build specific hand dataset and conduct more experiments to show how hand pyramid contributes for hand representation. We will discuss deeply on the results and analyze the impact of this pyramid on hand posture classification.

Keywords: Hand posture recognition · Hand pyramid structure · Kernel Descriptor

1 Introduction

Vision-based hand posture recognition plays an important role in natural human-machine interaction. Hand posture recognition takes a hand region image as a result of hand detection step and returns a label of hand posture. Challenges to vision-based hand posture recognition are the following: (i) vision-based hand posture recognition is affected by changes in lighting condition, cluttered backgrounds, and changes in scale; (ii) hand is a deformable object; there exist a considerable number of mutually similar hand postures; (iii) applications using hand posture generally require real-time, user-independent recognition.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_15

A number of hand recognition methods have been proposed to address these challenges [4,10,17,18]. These methods can be divided into two main categories depending how the hand is represented: explicit or implicit. Implicit representation relies on visual features which are computed directly from pixel values or reflect the relationship between pixels or regions. Simplest features are raw values of pixels. In this case, the image is often shaped into a 1D vector as the feature vector [12,13]. The using of raw pixel values makes high dimensionality of feature vector. Principal Component Analysis (PCA) is usually used [12,23] for dimension reduction. Instead of using directly pixel values, some other extract visual features that reflect relationship between pixels or regions. One of the popular features is extracted from Gabor filter [9]. Besides, many researchers also have employed Haar-like features [16,21,22]. Some other proposed to represent hands by orientation histogram [8], local binary pattern (LBP) [7] or SIFT [6] features.

The implicit hand representation methods do not reflect the structure of the hand. These methods hence do not take advantages of the specific structure of the hand. Some method using SIFT feature obtains good results on specific datasets that can provide a rich set of key points. However, when the resolution of hand image is low, the extracted set of key points is poor. In this situation, the method using SIFT feature will obtain a bad performance.

Explicit representation of hand bases on *intuitive* features. The methods belonging to this category often require good hand segmentation results to extract hand shape features (for example edges, contours) or topographical features such as fingers. Shape and topographical features are good for hand posture representation if we can segment well the hand region from the image [3,11,16,19,20]. However, hand segmentation is still a challenge in the real environment. Hand representation methods belonging to explicit presentation approach are intuitive and easy to understand for hand posture representation. However, these methods require a good hand segmentation result such as a clear hand contour. There was not a good combination between implicit and explicit representation. Some method used both of implicit and explicit features. Nevertheless, to use explicit features, we still need a good segmentation.

In our previous work [14], we proposed a method for hand presentation that is a flexible combination of implicit and explicit representation approaches. This hand representation method bases on kernel descriptor (KDES) [1]. We have made three improvements of KDES to make our descriptor more robust to orientation, scale change. One of the improvements was that we introduced a new pyramid structure which is intuitively more suitable to hand structure. However, we did not evaluate the impact of hand pyramid individually. In this paper, we build specific hand dataset and conduct more experiments to show how hand pyramid contributes for hand representation. We will discuss deeply on the results and analyze the impact of this pyramid on hand posture classification.

The remainder of the paper is organized as follows. The Sect. 2 reviews our method for hand posture recognition using the kernel method with hand pyramid structure. We then give some discussions and an evaluation of hand pyramid structure in the Sect. 3. The conclusions and directions for future work are given in the Sect. 4.

2 Hand Representation Using Hand Pyramid Structure

In [14], hand representation takes a hand region image (from now on called *image*, for short) as input and returns a descriptor of the hand candidate. It is composed of three sub-steps: Pixel-level feature extraction, Patch-level feature extraction, Image-level feature extraction.

2.1 Extraction of Pixel-Level Features

At pixel level, a normalized gradient vector is computed for each pixel of the image. The normalized gradient vector at a pixel z is defined by its magnitude $m(z)$ and normalized orientation $\omega(z) = \theta(z) - \overline{\theta}(P)$, where $\theta(z)$ is orientation of gradient vector at the pixel z, and $\overline{\theta}(P)$ is the dominant orientation of the patch P that is the vector sum of all the gradient vectors in the patch. This normalization will make patch-level features invariant to rotation. In practice, the normalized orientation of a gradient vector will be:

$$\widetilde{\omega}(z) = [sin(\omega(z))\ cos(\omega(z))] \tag{1}$$

2.2 Extraction of Patch-Level Features

Patch-level features are computed based on the idea of the kernel method. Derived from a match kernel representing the similarity of two patches, we can extract the feature vector for the patch using an approximate patch-level feature map, given a designed patch level match kernel function.

The gradient match kernel is constructed from three kernels that are gradient magnitude kernel $k_{\widetilde{m}}$, orientation kernel k_o and position kernel k_p. Gradient match kernel is defined as follows:

$$K_{gradient}(P, Q) = \sum_{z \in P} \sum_{z' \in Q} k_{\widetilde{m}}(z, z') k_o(\widetilde{\omega}(z), \widetilde{\omega}(z')) k_p(z, z') \tag{2}$$

where P and Q are patches of two different images that we need to measure the similarity. z and z' denote the 2D position of a pixel in the image patch P and Q.

The gradient magnitude kernel $k_{\widetilde{m}}$ is defined as:

$$k_{\widetilde{m}}(z, z') = \widetilde{m}(z)\widetilde{m}(z') \tag{3}$$

Where the normalized gradient magnitude $\widetilde{m}(z)$ is defined as:

$$\widetilde{m}(z) = \frac{m(z)}{\sqrt{\sum_{z \in P} m(z)^2 + \epsilon_g}} \tag{4}$$

where ϵ_g is a small constant. $m(z)$ is magnitude of the image gradient at a pixel z. The gradient magnitude kernel $k_{\widetilde{m}}$ is conspicuously a positive definite kernel.

Both the orientation kernel k_o and the position kernel k_p are Gaussian kernels which is of the form:

$$k(x, x') = \exp(-\gamma \|x - x'\|^2) \tag{5}$$

The factor γ will be defined individually for k_o and k_p that are denoted by γ_o and γ_p respectively.

Given the definition of match kernel, the approximate feature over image patch P is constructed as:

$$\overline{F}_{gradient}(P) = \sum_{z \in P} \widetilde{m}(z)\phi_o(\widetilde{\omega}(z)) \otimes \phi_p(z) \tag{6}$$

where $\otimes$ is the Kronecker product, $\phi_o(\widetilde{\omega}(z))$ and $\phi_p(z)$ are approximate feature maps for the kernel k_o and k_p, respectively. Given a match kernel function $k(x, y)$, the feature map $\varphi(.)$ for the kernel $k(x, y)$ is a function mapping a vector x into a feature space so as $k(x, y) = \varphi(x)^\top \varphi(y)$. Suppose that we have a set of basis vectors $B = \{\varphi(v_i)\}_{i=1}^{D}$, the approximation of feature map $\varphi(x)$ will be $\phi(x) = Gk_B(x)$, where G is defined by: $G^\top G = K_{BB}^{-1}$ and K_{BB} is $D \times D$ matrix with $\{K_{BB}\}_{ij} = k(v_i, v_j)$. k_B is a $D \times 1$ vector with $\{k_B\}_i = k(x, v_i)$.

The Kronecker product causes high dimension of the feature vector $\overline{F}_{gradient}$ (P). To reduce the dimension of $\overline{F}_{gradient}$, the kernel principal component analysis is applied into the joint basis vectors $\{\varphi_o(x_i) \otimes \varphi_p(y_j)\}_{i=1..d_o, j=1..d_p}$. Let t-*th* component ${\alpha^t}_{ij}$ is learned through kernel principal component analysis, following [1], the resulting gradient kernel descriptor for match kernel in (2) has the form:

$$\widetilde{F}^t_{gradient}(P) = \sum_{i=1}^{d_o} \sum_{j=1}^{d_p} \alpha^t_{ij} \sum_{z \in P} \widetilde{m}(z) k_o(\widetilde{\omega}(z), x_i) k_p(z, y_j) \tag{7}$$

2.3 Extraction of Image-Level Features Using Hand Pyramid Structure

Once patch-level features are computed for each patch, the remaining work is computing a feature vector representing the whole image. In [2], the authors used a spatial pyramid structure by dividing the image into cells using horizontal and vertical lines at several layers (Fig. 1(a)). This structure is generic, therefore does not take into account the specific shape of objects. In our work, as the hand is an object with a specific structure, we proposed a new pyramid structure specifically for the hand. In the following, we present in detail each step to build the final descriptor of the image.

Design a Hand Specific Pyramid Structure for Patch-Level Features Pooling: Fig. 1(b) shows the proposed hand pyramid structure. The main idea is to exploit characteristics of hand postures. Let the hand posture image have a size of $w \times h$. We observe that the regions at the image corners often do not contain information. For this reason, we only consider the area inside the inscribed

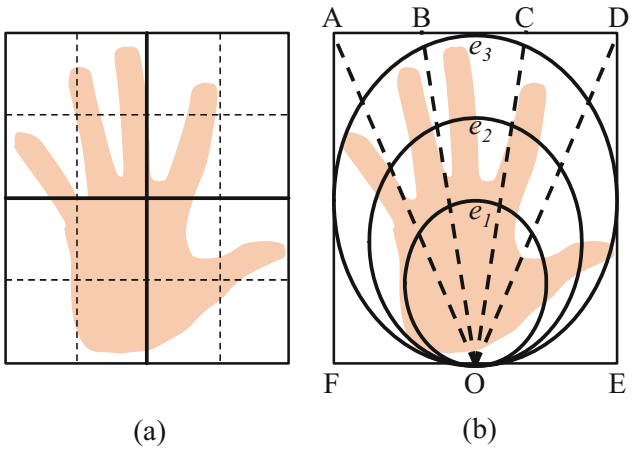


Fig. 1. (a) General spatial pyramid structure used in [2]. (b) The proposed hand pyramid structure.

ellipse of the hand image rectangle bounding box (e_3). The lines along the fingers converge at the lowest center point of the palm, near the wrist (O). Based on the structure of the hand, the ellipses (e_1, e_2, e_3) and the lines (OA, OB, OC, OD) are used to divide the hand region into parts that contain different components of the hand such as palm and fingers where $AB = BC = CD$. The detail of designed structure is described as: O is the midpoint of $FE (OF = OE)$. The ellipse e_1 is the inscribed ellipse of the rectangle that has a size of $(\frac{1}{2}w \times \frac{1}{2}h)$. The line FE is a tangent line of the ellipse e_1. The contact between the line FE and the ellipse e_1 is O. The ellipses are axis-aligned. In the similarity, the ellipsis e_2 is the inscribed ellipsis of the rectangle that has size of $(\frac{3}{4}w \times \frac{3}{4}h)$.

In a layer, we define a cell as being a full region limited by these ellipses and lines. In our work, the hand pyramid structure has 3 layers, (see Fig. 2).

- Layer 1: This layer contains only one cell defined by the biggest inscribed ellipse e_3.
- Layer 2: In [1], this layer has four rectangular cells. Unlike this, we create eight cells: three cells created from 3 ellipses and five cells created from the intersection of four lines with the biggest ellipse.
- Layer 3: This layer has 15 cells generated from the intersection between lines and three ellipses.

We can see that the hand pyramid gives a suitable representation for upright frontal hand postures in that the difference between postures is the configuration of the fingers (open or closed). In addition, the proposed hand representation has another advantage that is the image level feature is more invariant to slightly rotation/finger distortion. The reason of this ability is that the feature vector of a cell in pyramid is computed based on the set of feature vectors of patches belonging to the cell without concerning the locations of patches in the cell.

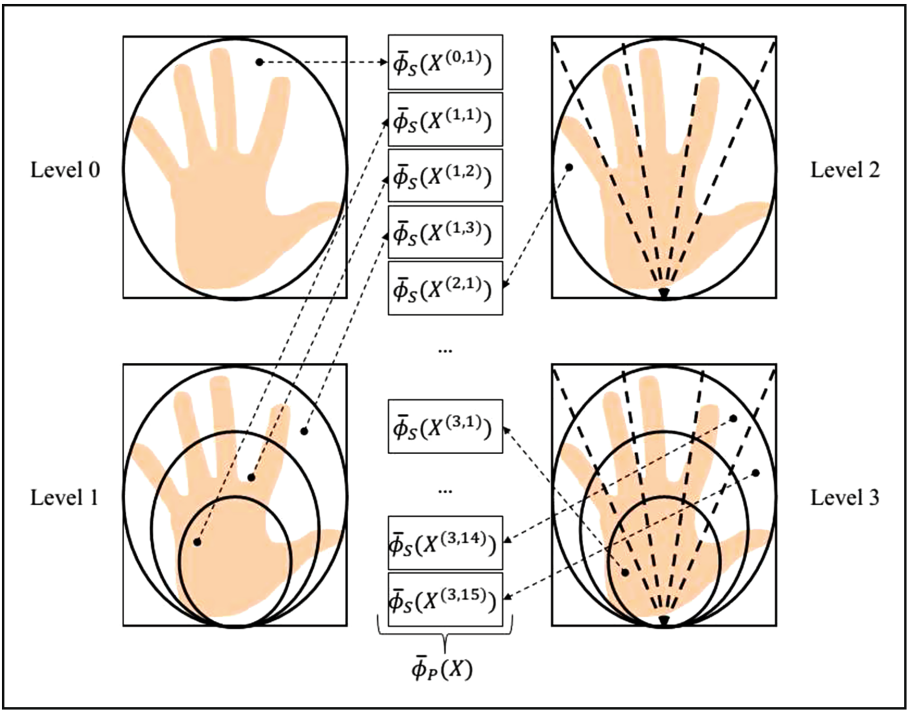


Fig. 2. Construction of image-level feature concatenating feature vectors of cells in layers of hand pyramid structure.

Moreover, the patch level feature is invariant to rotation. In case of our constraints (upright frontal hand postures with slightly rotation/finger distortion), the hand pyramid with normalized orientation gradient in the patch is suitable for hand representation, see Fig. 3(a,b). When hands strongly rotate or fingers heavily distort that make the same finger (or part of hand) in two hand images belongs to different cells in pyramid, above advantages of proposed hand presentation are not shown, see Fig. 3(c,d). In the case of strong rotation of hand, we could normalize hand image before applying hand pyramid.

Create the Final Descriptor of the Whole Image: To create the final descriptor of the whole image, we firstly use efficient match kernels (EMK) proposed in [1] to compute the feature vector for each cell of the hand pyramid structure, and then concatenate them into a final descriptor. Let C be a cell that has a set of patch-level features $X = \{x_1, ..., x_p\}$ then the feature map on this set of vectors is defined as:

$$\overline{\phi}_S(X) = \frac{1}{|X|} \sum_{x \in X} \phi(x) \tag{8}$$

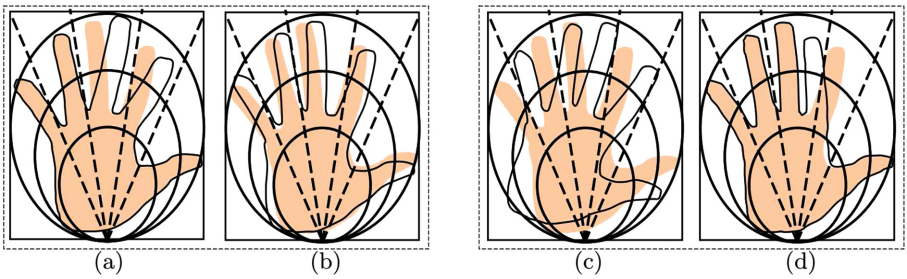


Fig. 3. (a) Slight finger distortion and (b) slightly hand rotation that do not make the same fingers belong to different cells; (c) heavy finger distortion and (d) strong hand rotation that make the same fingers belong to different cells.

where $\phi(x)$ is approximate feature maps for the kernel $k(x, y)$. The feature vector on the set of patches, $\overline{\phi}_S(X)$, is extracted explicitly.

Given an image, let L be the number of spatial layers to be considered. In this case $L = 3$. The number of cells in layer l-th is (n_l). $X(l, t)$ is set of patch-level features falling within the spatial cell (l, t) (cell t-th in the l-th level). A patch is fallen in a cell when its centroid belongs to the cell. The feature map on the pyramid structure is:

$$\overline{\phi}_P(X) = [w^{(1)}\overline{\phi}_S(X^{(1,1)}); ...; w^{(l)}\overline{\phi}_S(X^{(l,t)}); ...; w^{(L)}\overline{\phi}_S(X^{(L,n_L)})] \quad (9)$$

In (9), $w^{(l)} = \frac{\frac{1}{n_l}}{\sum_{l=1}^{L}\frac{1}{n_l}}$ is the weight associated with level l.

Figure 2 shows image-level feature extraction on the proposed hand pyramid structure. Until now, we obtain the final representation of the whole image, which we call image-level feature vector. This vector will be the input of a Multiclass SVM for training and testing.

3 Experimental Evaluation

3.1 Dataset

In order to evaluate the impact of hand pyramid structure on hand representation, we build a dataset that is a subset of L3i-MICA dataset [15]. This subset contains 13 upright frontal hand postures in that the difference between postures is the configuration of the fingers (open or closed), Fig. 4. In this dataset, hand posture images are captured in complex natural backgrounds. The hand regions are segmented manually. The number of training samples is 2859, and the number of testing samples is 2924. With this dataset, the accuracy we obtain in case of hand pyramid structure is 88.4 %, while the accuracy in case of general pyramid structure is 86.4 %. Hand pyramid structure makes the performance of hand posture recognition better (the accuracy increases 2 %). Thus we can remark that hand pyramid structure is suitable for upright frontal hand postures with open and closed fingers.

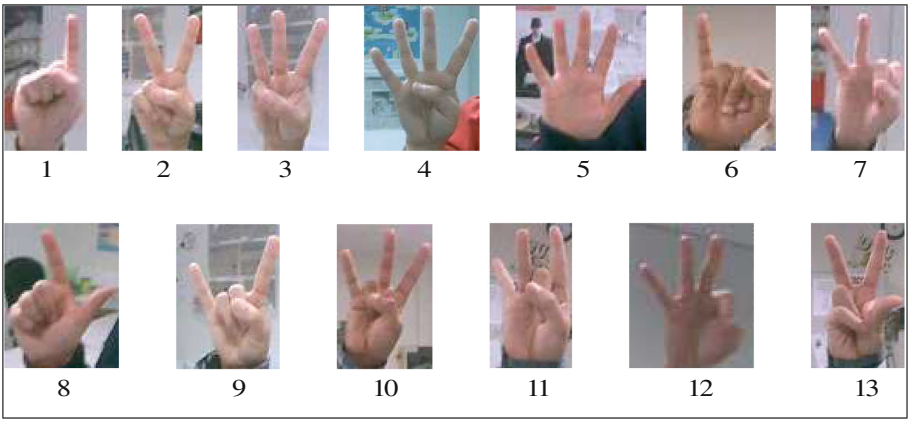


Fig. 4. Subset of L3i-MICA dataset.

4 Conclusions and Future Works

In this paper, we studied deeply the hand pyramid structure for hand representation in kernel method. We demonstrate empirically that our proposed hand pyramid is more suitable than general pyramid for hand posture recognition. This hand pyramid structure can be applied to detect open/closed fingers. In the future, we will evaluate hand pyramid structure with a more challenging dataset as well as a human-machine interaction using hand gestures.

Acknowledgement. This research is funded by Hanoi University of Science Technology under grant number T2016-PC-189.

References

1. Bo, L., et al.: Kernel descriptors for visual recognition. In: NIPS, pp. 1–9 (2010)
2. Bo, L., Sminchisescu, C.: Efficient match kernel between sets of features for visual recognition. In: NIPS, pp. 1–9 (2009)
3. Boughnim, N., Marot, J.: Hand posture recognition using jointly optical flow and dimensionality reduction. EURASIP J. Adv. Sig. Process. **2013**, 1–22 (2013)
4. Chaudhary, A., et al.: Intelligent approaches to interact with machines using hand gesture recognition in natural way: a survey. IJCSES **2**(1), 122–133 (2011)
5. Choi, J., Seo, B.: Robust hand detection for augmented reality interface. In: Proceedings of the 8th International Conference on Virtual Reality Continuum and its Applications in Industry, pp. 319–322 (2009)
6. Dardas, N.H., Georganas, N.D.: Real-time hand gesture detection and recognition using bag-of-features and support vector machine techniques. IEEE Trans. Instrum. Meas. **60**(11), 3592–3607 (2011)
7. Francke, H., Ruiz-del-Solar, J., Verschae, R.: Real-time hand gesture detection and recognition using boosted classifiers and active learning. In: Mery, D., Rueda, L. (eds.) PSIVT 2007. LNCS, vol. 4872, pp. 533–547. Springer, Heidelberg (2007). doi:10.1007/978-3-540-77129-6_47

8. Freeman, W.T., Roth, M.: Orientation histograms for hand gesture recognition. In: International Workshop on Automatic Face and Gesture Recognition, pp. 296–301 (1995)
9. Gupta, S., et al.: Static hand gesture recognition using local gabor filter. Procedia Eng. **41**, 827–832 (2012). Iris
10. Hasan, M.M., Mishra, P.K.: Hand gesture modeling and recognition using geometric features : a review. CJIPCV **3**, 1 (2012)
11. Kaufmann, B., Louchet, J., Lutton, E.: Hand posture recognition using real-time artificial evolution. In: Chio, C., Cagnoni, S., Cotta, C., Ebner, M., Ekárt, A., Esparcia-Alcazar, A.I., Goh, C.-K., Merelo, J.J., Neri, F., Preuß, M., Togelius, J., Yannakakis, G.N. (eds.) EvoApplications 2010. LNCS, vol. 6024, pp. 251–260. Springer, Heidelberg (2010). doi:10.1007/978-3-642-12239-2_26
12. Lockton, R., Fitzgibbon, A.W.: Real-time gesture recognition using deterministic boosting. In: BMVC, pp. 817–826 (2002)
13. Marcel, S., Bernier, O.: Hand posture recognition in a body-face centered space. In: Braffort, A., Gherbi, R., Gibet, S., Teil, D., Richardson, J. (eds.) GW 1999. LNCS (LNAI), vol. 1739, pp. 97–100. Springer, Heidelberg (1999). doi:10.1007/3-540-46616-9_9
14. Nguyen, V.-T., et al.: A new hand representation based on kernels for hand posture recognition. In: The 11th IEEE International Conference on Automatic Face and Gesture Recognition (FG 2015) (2015)
15. Nguyen, V.-T.: Visual interpretation of hand postures for human-machine interaction (2015)
16. Ong, E.-J., Bowden, R.: A boosted classifier tree for hand shape detection. In: Sixth IEEE International Conference on Face and Gesture Recognition, pp. 889–894 (2004)
17. Pavlovic, V.I., et al.: Visual interpretation of hand gestures for human-computer interaction: a review. TPAMI **19**(7), 677–695 (1997)
18. Rautaray, S.S., Agrawal, A.: Vision based hand gesture recognition for human computer interaction: a survey. AIR **43**, 1–54 (2012)
19. Sgouropoulos, K., et al.: A dynamic gesture and posture recognition system. J. Intell. Robot. Syst. **76**(2), 283–296 (2013)
20. Stergiopoulou, E., Papamarkos, N.: Hand gesture recognition using a neural network shape fitting technique. EAAI **22**(8), 1141–1158 (2009)
21. Tran, T-T-H Nguyen, T-T-M : Invariant lighting hand posture classification. In: IEEE International Conference on Progress in Informatics and Computing (PIC), pp. 827–831 (2010)
22. Wachs, J., et al.: A real-time hand gesture system based on evolutionary search. In: GECCO2005: Tenth Genetic and Evolutionary Computation Conference, Washington, DC, USA (2005)
23. Wu, Y., Huang, T.S.: View-independent recognition of hand postures. In: Computer Vision and Pattern Recognition, pp. 88–94 (2000)

An Evolutionary-Based Term Reduction Approach to Bilingual Clustering of Malay-English Corpora

Rayner Alfred(✉), Leow Ching Leong, and Joe Henry Obit

Faculty of Computing and Informatics, Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia
{ralfred,joehenry}@ums.edu.my, dragon_july14@hotmail.com

Abstract. The document clustering process groups the unstructured text documents into a predefined set of clusters in order to provide more information to the users. There are many studies conducted in clustering monolingual documents. With the enrichment of current technologies, the study of bilingual clustering would not be a problem. However clustering bilingual document is still facing the same problem faced by a monolingual document clustering which is the "curse of dimensionality". Hence, this encourages the study of term reduction technique in clustering bilingual documents. The objective in this study is to study the effects of reducing terms considered in clustering bilingual corpus in parallel for English and Malay documents. In this study, a genetic algorithm (GA) is used in order to reduce the number of feature selected. A single-point crossover with a crossover rate of 0.8 is used. Not only that, this study also assesses the effects of applying different mutation rate (e.g., 0.1 and 0.01) in selecting the number of features used in clustering bilingual documents. The result shows that the implementation of GA does improve the clustering mapping compared to the initial clustering mapping. Not only that, this study also discovers that GA with a mutation rate of 0.01 produces the best parallel clustering mapping results compared to GA with a mutation rate of 0.1.

Keywords: Clustering bilingual documents · Hierarchical agglomerative clustering · Evolutionary algorithm · Genetic algorithm · Corpus

1 Introduction

Document clustering is an unsupervised machine learning technique in text mining. The purpose of document clustering is to group documents into clusters. The documents in a cluster have similar characteristics shared among them. The most well-known clustering technique in document clustering is agglomerative and partitional. There are advantage and disadvantage in each technique. Partitional clustering is best applied in large dataset whereas hierarchical clustering time complexity is higher compared to partitional clustering. Hierarchical clustering produces more accurate results than partitional clustering.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_16

However, basic clustering technique is insufficient to produce a good clustering quality in document clustering. One of the problems is due to the high dimensionality of documents. High dimensionality is also known as *curse of dimensionality* [2]. News articles may consist of many unique terms. Having a huge number of terms will definitely increase the complexity of the computation and it indirectly affects the clustering results. There are lots of studies of term reduction in document clustering [11,14]. However, a clustering bilingual language for English and Malay corpus is still at early stage. This study tends to increase clustering mapping by optimizing the terms for Malay news articles and English news articles. There are several advantages on optimizing the terms on bilingual clustering. Term reduction in bilingual clustering reduces twice the time complexity. Not only that, the extracted terms can be shared on both languages which will provide more information to users. It is anticipated that by improving the clustering results, it will also improve the mapping of Malay and English clusters obtained from the bilingual clustering. Hence, by increasing the mapping percentage for the bilingual clusters, a more robust clustering algorithm can be developed for clustering bilingual documents. In this study, a genetic algorithm (GA) is proposed to be implemented in order to determine the set of terms that can be used in clustering bilingual documents with more effective.

This paper is divided into several sections. Section 2 discusses the implementation of genetic algorithm in document clustering. Not only that, it will also discuss related works implemented in clustering English news articles and Malay news articles. Section 3 explains the flow of the proposed work. In this section, each step in the proposed framework will be explained. Not only is that, the techniques applied for English document clustering and Malay document clustering are also elaborated. Section 4 shows the results of this experiment and the discussion on the experimental results. This result compares the clustering mapping percentage before implementing genetic algorithm and also the effect on the clustering mapping percentage after applying genetic algorithm. Section 5 concludes this paper with the discussion on future works that can be performed in order to improve the information retrieval for bilingual clustering which is in English and Malay.

2 Related Works

There are several term reduction techniques that had been applied in document clustering such as Latent Semantic Indexing [3], Random Projection [4], principal component analysis [5], nonnegative matrix factorization [6] and Independent Component Analysis [7]. Most of these techniques focus on arranging the terms based on the term weight in the constructed matrix. There is a possibility that low frequency terms do provide some hidden information. For example, a document stating news related to "car crash" may not have a high frequency value for the word "accident". Therefore, this study proposes the implementation of genetic algorithm in order to optimize the terms that can be shared among the news articles and also for both languages cluster.

Genetic algorithm (GA) is one of the famous evolutionary algorithms and it was proposed by John Holland during the 70s with the inspiration of survival of the fittest which is Darwin theory [8]. The concept of genetic algorithm is that the gene evolves to adapt to current environment. There are two areas in which a GA can be used for term reduction which are feature selection and feature weighting. GA had been applied in several studies for feature selection. One of the studies is done by Jarmulak and Craw [9]. In this study, GA had been used in a wrapper feature selection method in order to select the features and weighting the features. The result in this study shows that the implementation of GA as a wrapper method has improved the classification task for both selection and weighting. There are several studies that apply GA in bilingual clustering as feature weighting. One of the studies is on clustering Bulgarian-English corpus [10]. The implementation of GA in this study is to adjust the weight of the terms in order to increase the mapping of the clustering results. This study applies roulette wheel selection, single-point crossover and uniform mutation. The rates that will be used in this study are 0.25 and 0.01 for the crossover and mutation rates respectively.

Another study for feature weighting is proposed by Alfred et al. [11]. In this study, a GA based algorithm is applied that adjusts the weight of Term Frequency-Inverse Document Frequency (TF-IDF) by implementing a genetic algorithm. The weighting schemes used in this study include 0.5, 1.0 and 1.5. The genetic operators used are single-point crossover operation with probability rate of 0.25 and uniform mutation rate of 0.01 [10]. The study shows that the implementation of GA in HAC with average linkage show the best performance in clustering mapping between English news articles and Malay news articles. This study investigated the effects of feature weighting implemented in a bilingual clustering for English news articles and Malay news articles on the percentage of clustering mapping between the two cluster results obtained. However, there is no study conducted to investigate the effects of feature selection implemented in a bilingual clustering for English news articles and Malay news articles on the percentage of clustering mapping between the two cluster results obtained. Hence, in this paper, an evolutionary-based term reduction approach to bilingual clustering of Malay-English corpora is proposed and evaluated.

3 Experimental Setup

The flow of work is as shown in Fig. 1. There are 500 Malay news articles and 500 English news articles that will be used in this study. These news articles are retrieved from Bernama archive and theStar website. The news articles are retrieved manually to make sure that the contents of both bilingual articles are the same but they are written in different languages. Before the news articles can be clustered, there are several data pre-processing methods that should be applied in the news articles. There are lots of unique terms in news articles. However, not all the terms are considered useful in document clustering. Such terms are considered as noises that affect the quality of clustering results.

The elimination of such noises does help in reducing computational times in irrelevant terms. Other than that, there is also study proven that the data pre-processing does helps in improving document clustering results [12]. The data pre-processing helps to increase the efficiency of clustering results and improved on the information retrieval. The data pre-processing that involved in this study are *stopword elimination*, *stemming* and *named-entity recognition elimination*. However, there is no single data pre-processing algorithm that is applicable to all languages. Hence these data pre-processing techniques implementing in English news articles and data pre-processing techniques for Malay news articles are different.

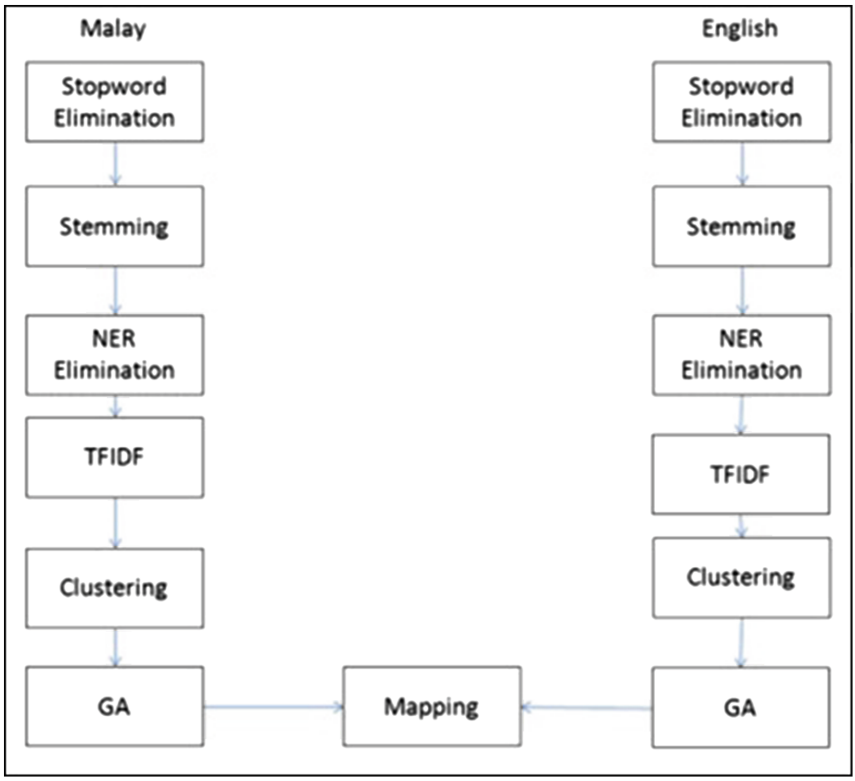


Fig. 1. Flow work of parallel clustering with term reduction.

The first data pre-processing process is the *stop words* elimination. *Stop words* are words that usually exist in documents and do not represent any meaning in the documents (e.g., the words *the*, *a*, *is* for English and *telah*, *itu*, *ke* for Malay). *Stop words* usually are in high frequency. The elimination of *stop words* will reduce the number of terms in a document. Eliminating the *stop words* will

reduce the computational complexity and also time complexity [13]. There are two sets of list of *stop words* used in this experiment.

The next process in data pre-processing is *stemming*. *Stemming* is a process of transforming the words to its root word [14]. In this work, there are two types of stemming algorithms used for processing bilingual documents. First, the Porter stemmer is applied to transform English words to its root words [8]. As for the Malay stemming algorithm, a rule-based Malay stemming algorithm will be applied [16].

The next process is eliminating the named-entity with the help of the Named-Entity Recognition (NER). This NER process helps users to produce more meaningful corpus by identifying proper names [11,18]. In this study, the Stanford NER is applied in order to recognize English NEs [13,16]. As for Malay NER, a rule-based Malay NER had been applying in this study. There are not many NER systems available for Malay language. This Malay NER apply rule-based such as preposition, location prefix, person first name, person middle name, organization prefix, organization suffix and list of gazetteers to detect the entities [14].

After data pre-processing, the terms that still available in the documents are considered important for document clustering. The Term Frequency-Inverse Document Frequency (TF-IDF) [22] will be applied to weight each term that exists in the corpus. The equation of TF-IDF is as shown below,

$$tfidf(d,t) = tf(d,t) * log\frac{|D|}{df(t)} \tag{1}$$

where *df(t)* is the frequency of documents that term t appears. Then clustering will be implemented to the news articles. In this study, a hierarchical agglomerative clustering (HAC) [10] is applied. In this study, HAC coupled with average linkage had been implemented. The HAC coupled with average linkage measure the distance between clusters by calculate the average distance between documents of the compared clusters C_i and C_j, as shown below,

$$proximity(C_i, C_j) = \frac{\sum_{x \in C_i, y \in C_j} proximity(x,y)}{m_i * m_j} \tag{2}$$

where m_i and m_j are the number of documents in each cluster. The average distance between clusters changed on each merging of the clusters as it depends on the number of documents in the cluster.

Extend Jaccard coefficient had been used to measure the proximity matrix between the documents. It compares the sum weight of shared terms to the sum weight of terms that are present in either of the two documents but are not the shared terms [15] as shown below, where x and y is the vectors of the respective documents and $||x|| = sqrt(x * x)$ and $||y|| = sqrt(y * y)$.

$$ExtendJaccard = \frac{x * y}{||x||^2 + ||y||^2 - x * y} \tag{3}$$

In this work, a chromosome will represent a set of terms that will be selected and considered when clustering bilingual documents in parallel. For example,

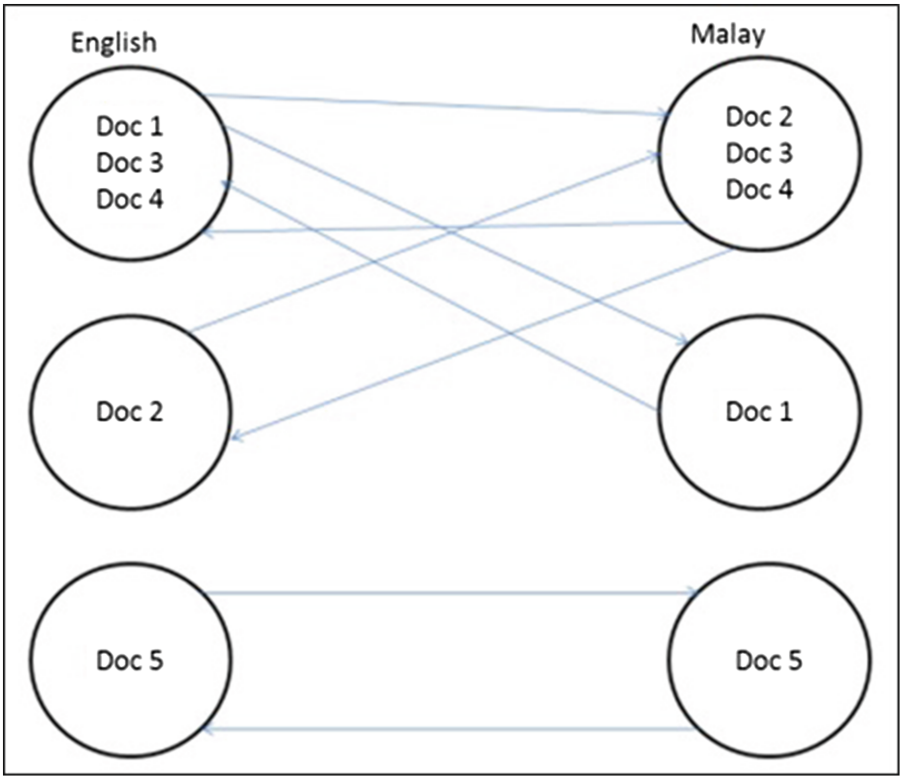


Fig. 2. Flow work of parallel clustering with term reduction.

a chromosome 01011 shows that term2, term4 and term5 are used for clustering. This study will not investigate the effects of varying the rates of crossover on the clustering results. A generation of 20 chromosomes is generated randomly and there are a total of 50 generations generated in this experiment. A basic crossover and mutation concept will be applied in this study with crossover rate of 0.8 and with mutation rate of 0.01 and 0.1 are used.

Parallel clustering mapping percentage is used as the fitness function in the GA as the purpose of this study is to improve the cluster mapping between bilingual clustering. Figure 2 shows the example of mapping the clustering results between English clusters and Malay clusters. It consists of 5 articles and they are clustered into 3 clusters. For English articles, Cluster 1 consists of Doc 1, Doc 3 and Doc 4. Cluster 2 contains Doc 2 and Cluster 3 contains Doc 5. There are slightly different on Malay clustering result compared to English clustering result. Doc 2, Doc 3 and Doc 4 are clustered in Cluster 1. Doc 1 is grouped into Cluster 2 and Doc 5 is grouped into the last cluster.

Mapping percentage needed to be calculated from bi-directions in order to make sure that both clustering are parallel. Other than that, if the mapping

percentage is based on one direction flow only, then 100 % will be achieved when English Cluster 2 is mapped to Malay Cluster 1. However, the number of matched articles is less compared to English Cluster 1. Hence, this encourages that mapping percentage should be calculated based on a bi-direction flow. Since this experiment focuses on clustering bilingual documents, the percentage of mapping bilingual clusters is used to determine the bilingual clustering quality. Percentages of mapping bilingual clusters can be used to show the similarity of bilingual clustering results. Not only that, the parallel clustering results can be used to show that several terms can be shared by both languages in clustering bilingual documents and reveal more hidden information. The formula used for mapping the bilingual clusters between English cluster and Malay cluster is as followed,

$$Mapping(a,d) = \sum_{1 \le i \le k, 1 \le j \le k} \frac{\frac{|C_i(a) \cap C_j(d)|}{C_i(a)} * \frac{|C_i(a) \cap C_j(d)|}{C_j(d)}}{k} \tag{4}$$

where $|C_i(a)|$ is the number of documents in English cluster and $|C_j(d)|$ is the number of Malay documents in Malay clusters. Then, $|C_i(a) \cap C_j(d)|$ refers to the number of documents in English cluster that can be mapped to Malay documents or vice versa.

4 Results and Discussion

Table 1 shows the clustering mapping percentage for various cluster size. The cluster's size is 5, 10, 15, 20 and 25. The table shows the cluster mapping percentage for initial cluster mapping without the implementation of GA. The third column of the table shows cluster mapping results for GA with mutation rate of 0.1. The last column of the table shows clustering mapping percentages for the implementation of GA with mutation rate of 0.01. Based on the results, it shows that the mapping percentage is quite low for cluster size 5. This shows that there are fewer clusters that shared the same news articles between English cluster and Malay cluster. The mapping percentage increases when the cluster's size is 10. Similar mapping percentages can be obtained when the cluster size are 15, 20 and 25 which are roughly 51 %.

The results show that the GA based clustering technique is able to improve the results of clustering mapping compared to the initial results obtained.

Table 1. Cluster mapping percentage for various cluster sizes

Cluster size	Initial cluster mapping	GA with mutation rate 0.1	GA with mutation rate 0.01
5	38.88 %	79.66 %	79.37 %
10	47.64 %	53.65 %	67.96 %
15	51.68 %	51.89 %	54.31 %
20	51.3 %	55.34 %	52.73 %
25	51.25 %	54.77 %	55.10 %

The mapping percentage obtained from the initial results for all cluster sizes are lower compared to cluster mapping percentages obtained when using the GA based bilingual clustering. The best mapping percentage is obtained when the cluster size is 5 coupled with mutation rates of 0.1. Based on the table shown, the results show that the mapping percentages of bilingual clusters can be improved by optimizing the terms selection process. Even though there are some improvements on the clustering mapping, the mapping percentages decrease as the size of cluster increases. The percentage of clustering mapping with cluster size of 5 shows a great improvement. The highest improvement value of the percentage of clustering mapping is 67.96 % when the cluster's size is 10. When the clusters' sizes are 15, 20 and 25, the percentages of improvement obtained are just roughly 55 %. The highest cluster mapping percentage is highlighted in bold for each cluster size. The GA based clustering technique performs better on others cluster size when the mutation rate is 0.01. Hence, the GA based clustering technique with a mutation rate of 0.01 is better than 0.1 as it produces most of the highest clustering mapping percentages.

5 Conclusion

As a conclusion, the implementation of GA in clustering English news articles and Malay news articles does improve the clustering mapping between bilingual clustering. This also proven that the feature selection can be used to improve the bilingual clustering mapping. Not only that the results also show that, the GA with a mutation rate of 0.01 shows a better clustering mapping. Unfortunately, a perfect result still cannot be achieved. This shows that the parameter setting with 50 generations generated with 20 chromosomes produced in each generation are insufficient in order to produce the perfect score. In future works, more generations and populations can be applied in the study. Not only that, the different type of crossover and crossover rates can also be applied in order to study the effects of different rates on the clustering bilingual documents. Instead of using a single-point crossover, a two-point crossover or uniform crossover which is suitable to deal with long length chromosome can be also applied. Not only that, the crossover rate used in this study is 0.8. Hence, more experiments can be conducted for other crossover rates (e.g., 0.9 and 1.0).

Acknowledgments. This work has been supported by the Long Term Research Grant Scheme (LRGS) project funded by the Ministry of Higher Education (MoHE), Malaysia under Grants No. LRGS/TD/2011/UiTM/ICT/04.

References

1. Tang, B., Shepherd, M., Heywood, M.I., Luo, X.: Comparing dimension reduction techniques for document clustering. In: Kégl, B., Lapalme, G. (eds.) AI 2005. LNCS (LNAI), vol. 3501, pp. 292–296. Springer, Heidelberg (2005). doi:10.1007/11424918_30
2. Alfred, R., Chan, C.J., Tahir, A., Obit, J.H.: Optimizing clusters alignment for bilingual Malay-English Corpora. J. Comput. Sci. **8**(12), 1970–1978 (2012)
3. Alfred, R., Paskaleva, E., Kazakov, D., Bartlett, M.: Hierarchical agglomerative clustering of English-Bulgarian parallel Corpora. In: Proceedings of International Conference of Recent Advances in Natural Languages Processing (2007)
4. Micheal, W.B., Susan, T.D., Gavin, W.O.B.: Using linear algebra for intelligent information retrieval. SIAM Rev. **37**(4), 573–595 (1995)
5. Bingham, E., Mannila, H.: Random projection in dimensionality reduction: applications to image and text data. In: Proceedings of ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, pp. 245–250 (2001)
6. Andrews, N.O., Fox, E.A.: Recent developments in document clustering. Technical report TR-07-35. Computer Science. Virginia Tech (2007)
7. Xu, W., Liu, X., Gong, Y.H.: Document clustering based on non-negative matrix factorization. In: Proceedings of the 26th Annual International ACM SIGIR Conference on Research and Development in Information Retrieval, pp. 267–273 (2003)
8. Hyvarinen, A., Oja, E.: Independent component analysis: algorithms and applications. Neural Netw. **13**(4), 411–430 (2000)
9. Holland, J.H.: Adaptation in Natural and Artificial Systems. MIT Press, Cambridge (1992)
10. Hussein, F., Ward, R., Kharma, N.: Genetic algorithms for feature selection and weighting, a review and study. In: 12th International Conference on Document Analysis and Recognition, p. 1240 (2001)
11. Alfred, R., Paskaleva, E., Kazakov, D., Bartlett, M.: Hierarchical agglomerative clustering for cross-language information retrieval. Int. J. Transl. **19**(1), 1–25 (2007)
12. Kumar, A.A., Chandrasekhar, S.: Text data pre-processing and dimensionality reduction techniques for document clustering. Int. J. Eng. Technol. (IJERT) **1**(5), 1–6 (2012)
13. El-Khiar, I.: Effects of stops words elimination for Arabic information retrieval: a comparative study. Int. J. Comput. Inf. Sci. **4**(3), 119–133 (2006)
14. Porter, M.: An algorithm for suffix stripping. Program (Autom. Libr. Inf. Syst.) **14**(3), 130–137 (2006)
15. Leong, L.C., Basri, S., Alfred, R.: Enhancing Malay stemming algorithm with background knowledge. In: Anthony, P., Ishizuka, M., Lukose, D. (eds.) PRICAI 2012. LNCS (LNAI), vol. 7458, pp. 753–758. Springer, Heidelberg (2012). doi:10.1007/978-3-642-32695-0_68
16. Montolva, S., Martinez, R., Casillas, A., Fresno, V.: Multilingual document clustering: an heuristic approach based on cognate named entities. In: Proceedings of COLING-ACL, pp. 1145–1152 (2006)
17. Montalvo, S., Fresno, V., Martinez, R.: NESM: a named entity based proximity measure for multilingual news clustering. Procesamiento del Lenguaje Nat. **48**, 81–88 (2012)

18. Finkel, J.R., Grenager, T., Manning, C.: Incorporating Non-local information into information extraction systems by gibbs sampling. In: Proceedings of the 43rd Annual Meeting of the Association for Computational Linguistics (ACL 2005), pp. 363–370 (2005)
19. Atdag, S., Labatut, V.: A comparison of named entity recognition tools applied to biographical texts. In: 2nd International Conference on Systems and Computer Science (ICSCS), pp. 228–233 (2013)
20. Alfred, R., Leow, C.L., Chin, K.O., Anthony, P.: Malay named entity recognition based on rule-based approach. IJMLC **3**(4), 300–306 (2014)
21. Manning, C.D., Raghavan, P., Schutze, H.: An Introduction to Information Retrieval, 1st edn. Cambridge University Press, Cambridge (2009)
22. Huang, A.: Similarity measures for text document clustering. In: Proceedings of the 6th New Zealand Computer Science Research Student Conference, NZCSRSC, pp. 49–56 (2008)

An Exploratory Study on Students' Performance Classification Using Hybrid of Decision Tree and Naïve Bayes Approaches

Yoong Yen Chuan(✉), Wahidah Husain, and Amirah Mohamed Shahiri

School of Computer Sciences, Universiti Sains Malaysia,
11800 USM, Penang, Malaysia
bearyoong0806@gmail.com, wahidah@usm.my, ams14_com044@student.usm.my
http://www.usm.my

Abstract. Students' performance prediction can give a prior approximate knowledge of the students' performance in future academic to the educators. However, it is not any easy task to perform prediction due to the poor identification of parameters and the lack of prediction techniques. In this paper, few parameters will be proposed and the most influenced parameters on students' performance will be identified using chi squared. The hybrid of Decision Tree and Naïve Bayes algorithms, NBTree will be used to classify the performance of new students. NBTree classifier undergoes the training and testing process using 10-folds cross validation technique and obtained the classification accuracy of 85.9 %, which is better than the accuracy of Decision Tree and Naïve Bayes classifiers which are having 63.7 % and 72.6 % respectively. The classified performance result can be used by the educators to improve the teaching and learning process by developing new teaching methods and new teaching styles.

1 Introduction

Students' performance prediction is the prediction of students' future academic achievement in term of GPA. It is usually used for predicting the performance of new intake students by the educators to have an early approximate knowledge of the students' future academic performance. Based on the predicted performances, educators can categorize the students in different category groups, such as good, average and poor. Thus, they can spend more time teaching and pay more attention helping those students that are categorized in poor group. Moreover with the prediction, the educators can come out with new teaching methods such as flipped classroom, massive open online course (MOOC) and etc. to assist the weak students and improve the teaching and learning process.

Data mining is the process of analyzing and extracting large set of data usually business market related from different perspective view and generate in into useful information. It is also one of the famous techniques used for prediction in the educational system, which is called educational data mining (EDM) [1].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_17

EDM uses computational approaches to analyze and understand educational data in order to study questions related to the students' educational problem. The development methods of EDM came from the application of data mining algorithm, machine learning and statistic to analyze collected students' data during teaching and learning process.

With data mining algorithm, it helps to classify the students' performance by analyzing the extracted data [2]. However, measuring students' performance is not an easy task as it hinges on various kinds of unknown factors such as students' demographic, family background, school area and etc. Thus, the objectives of this study are:

1. To identify the predictive parameters on students' performance.
2. To hybrid the Decision Tree and Naïve Bayes algorithms in order to design, implement and evaluate for students' performance classification using the identified parameters

The next section will focused on the literature review of Decision Tree and Naïve Bayes followed by the related work on parameters used for prediction. Then, research methodology is discussed in Sect. 3. Section 4 explains the result and evaluation of the NBTree classifier while the conclusion and future work are outlined in Sect. 5.

2 Literature Review

In this section, the classification in educational data mining is studied and the related work of parameters used for prediction is reviewed.

2.1 Classification in Educational Data Mining

Data mining is being widely used in the educational field for prediction, known as Educational Data Mining (EDM). EDM is used to predict students' performance by applying data mining algorithms which also known as machine learning algorithm. A machine learning algorithm is a class of algorithms which is data-driven, unlike normal algorithm, it is the data that "tells" what the "good answer" is [3].

For students' performance prediction, the most suitable data mining approach is classification. Data mining classification will organize the data into categories in order to extract useful information for efficient use. The classifiers under the classification task are using supervised learning characteristic by mapping the input variables into discrete categories. In this research, Decision Tree and Naïve Bayes classifiers will be studied to understand the strength and weakness of the algorithm.

Decision Tree Classifier. Decision Tree is a graphical representation of possible solutions to a decision based on certain conditions given. Decision Tree is commonly used for retrieving information for the purpose of decision-making. It is like a flowchart tree structure where it starts from a single box or normally called root node, and then branches off into various numbers of solution called leaf nodes according to different decision tree learning algorithm [4]. The final outcome for decision tree is based on the representation of each branch as it contains a possible scenario of decision and result.

There are two common issues for the construction of decision tree which are the growth of the tree to enable it to accurately categorize the training data set and the pruning stage [2]. This classifier is commonly used by most of the researchers due to its simplicity and comprehensibility to reveal small or large data set and predict the value [5]. However, it also has weaknesses as this classifier can be complex and timeconsuming in preparing when having a large set of data with many branches.

Naïve Bayes Classifier. Naïve Bayes classifier is a probabilistic classifier which based on Bayes' Theorem. It is called naïve because it simplifies the matter based on 2 assumptions [6]: it assumes that all attributes are independent of each other, and there are no hidden attributes that will affect the prediction. For example, a fruit may be treated as an orange if it is orange color, round and having approximately 3 in. diameter. Despite these features depend on each other or upon the presence of other features, all of these properties independently showing the probability that this fruit is an orange [7]. Naïve Bayes predicts the probability of different class based on various attributes. It is efficient for data classification despite its simplicity. However, Naïve Bayes classifier is poor at predicting the true class conditional probabilities [8].

2.2 Related Work of Parameters Used in Prediction

Attributes play an important role in predicting students' performance as it can influence the prediction result. When the most influential variables are known, the building process of the prediction model will become very easy [9]. There are many attributes that had been applied by previous researchers in their study to identify the most influenced parameters on estimating students' grade. Out of 20 researches' work, the list of commonly used parameters for predicting students' performance is shown in Fig. 1.

From the figure, the most frequently used attributes for predicting students' performance is cumulative grade point average (CGPA) or students' past performance. Fifteen out of twenty researchers have used CGPA as their main attributes to predict students' performance [5,6,10–22]. The performance indicator used by them is mostly class labels which classify the students' CGPA into few categories such as poor, fair and good. The reason for them to choose CGPA as their main parameter is that CGPA has a tangible value for future educational and career mobility [23]. Students' demographic is the second most frequently

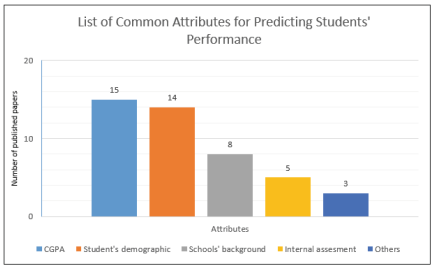


Fig. 1. List of common parameters used for predicting students' performance

used attributes for predicting students' performance. Students' demographic consists of students' age, gender, family background, such as parents' education, parents' occupation, family income and etc. Fourteen out of twenty researchers have chosen this parameter to measure students' performance [6,10,12,13,15,16,19–22,24–27]. Family background can also influence the students' performance in school as students spent most of their time at home besides going to school.

The third most frequently used attributes for predicting students' performance is schools' background. Schools' background attribute consists of the school environment, schools' area, type of school and etc. Eight out of twenty researches applied this attribute in predicting students' performance because school is the second place where students' spend most of their time [6,13,16,18,20,24–26].

Besides these three main constantly used parameters, internal assessment and other attributes such as socioeconomic, emotional skills and etc. are also being used by past researchers to predict students' performance. Internal assessment consists of quizzes, assignment mark, laboratory mark and etc. Based on the graph, five out of twenty researchers have chosen internal assessment attributes to estimate students' academic outcome [17,22,24,28,29]. On the other hand, only four out of twenty researchers will use other attributes to predict the students' result [11,12,21,27].

3 Methodology

There are four main steps for conducting this research work and it is shown in Fig. 2.

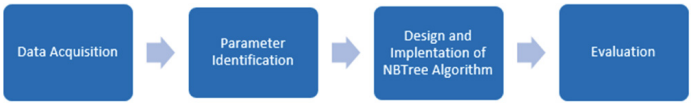


Fig. 2. Four main research methodology steps

3.1 Data Acquisition

The dataset needed for conducting this research will be collected through an online survey. Two types of students' data which are the graduated students' data and newly enrolled students' data were collected. A total of 17 attributes (1 output and 16 nominal variables) was collected from the online survey. The list of collected attributes and its respective values is shown in Table 1. The performance indicator used for the parameters will be class label using Good, Average and Poor categories. Moreover, the parameters are using multi-class classification where each variable are restricted to have only one label. The reason of using multi-class instead of multi-label is because the outcome of this research is to get the classification of students in term of performance category only, so that educators can group the students easily.

Table 1. List of collected attributes

Attributes	Possible values
Gender	{Male, Female}
Parents educational level	{Primary, Secondary, Tertiary}
Family size	{Small, Medium, Large}
Family income	{Low, Average, High}
Primary school performance	{Poor, Average, Good}
Secondary school performance	{Poor, Average, Good}
Attendance in class	{Poor, Average, Good}
Internal assessment	{Poor, Average, Good}
Social level	{Poor, Average, Good}
Study duration per day	{Short, Average, Long}
Attention time span in class	{Poor, Average, Good}
Sleeping duration per day	{Short, Average, Long}
Learning ability	{Low, Average, High}
Daily diet	{Poor, Average, Healthy}
Revision for test	{One day before, Three days before, One week before}
Relationship	{True, False}
Performance grade	{Poor, Average, Good}

The data obtained from the survey will be undergoes data cleaning through WEKA before processed. The cleaning process will eliminate the data with missing values, correcting inconsistent data, identifying outliers, as well as removing duplicated data. From WEKA choose filter → unsupervised attributes → Replace Missing Values and filter → unsupervised instances → Remove Duplicate to clean the dataset [25]. After the data cleaning process, a total of 150 graduated students' data ended up with 135 instances and 45 instances as the new students' data. These data are then used for training and testing the algorithm.

3.2 Parameter Identification

Predictive parameters used to predict students' performance were selected based on the comprehensive studies and analysis on articles and journals related to the research topic. The parameters were also selected based on the personalities of University Sains Malaysia students. Attribute selection techniques were used to find out the most relevant attributes in the data. Attribute selection will search through all the possible combinations of attributes in the data to find out which subset of attributes works best for prediction [25]. It is useful in reducing the dimensionality of the data to be processed by the classifier, reducing the execution time and improving the predictive accuracy [30]. In order to utilize attribute selection, two objects which are attribute evaluator and search method must be inserted. The evaluator acts as the quality metric to evaluate the predictive properties of an attribute.

Chi Squared Attributes Evaluator and Ranker search methods were selected to identify the best attributes among the 16 of them. Chi Squared will evaluate the attributes individually by measuring the chi-squared statistic with respect to the class and the correlation with it. This independent test method will identify the degree of association between variables and analyze the dependency of all attributes on the outcome attribute [31]. Ranker search method is used to rank the attributes by their individual evaluations according to the chosen quality metric and met the criterion. The Ranker search method performs the rank on which attributes should be obtain high or low rank based on the selected attributes in the given dataset. It provides a rating of the attributes orderly by their score to the evaluator [32]. The list of attributes and their ranked when using a full training set are shown in Table 2. The top 10 attributes with the highest rank will be selected because the ranked values of the attributes after the 10th are getting smaller and not significant enough to affect the performance of the prediction model. Thus, those parameters can be eliminated.

3.3 Design and Implementation of NBTree Algorithm

The proposed algorithm, NBTree consists of Naïve Bayesian classification and Decision Tree learning. It sorts the example of a leaf and builds a Naïve Bayes classifier on each leaf node of the built Decision Tree which integrates the advantage of both classifiers [33]. The algorithm uses Decision Tree to segment the training data where each segment of training data is represented by leaf node of the trees, and then builds a Naïve Bayes classifier on each segment [34]. The visualization of NBTree model is shown in Fig. 3.

With this approach, it makes the segmentation and leaf easy to be understood. Besides, the Naïve Bayes algorithm does not perform well with continuous valued attributes while Decision Tree does not produce good results when the data size is extremely large. Thus, this proposed hybrid algorithm, NBTree will overcome both the weaknesses of Decision Tree and Naïve Bayes classifiers. It also had advantage which is the simplicity of representing the learned knowledge.

Table 2. Highly potential variables

Attributes	Ranked values
Learning ability	37.82
Primary school performance	34.30
Sleeping duration per day	33.93
Attention time span in class	27.83
Internal assessment	27.54
Attendance in class	25.93
Study duration per day	24.01
Parents educational level	22.14
Daily diet	19.43
Secondary school performance	13.71
Family income	8.66
Social level	7.58
Relationship	5.47
Family size	5.33
Revision for test	5.01
Gender	0.17

3.4 Evaluation

The performance measures of NBTree classifier in term of accuracy, sensitivity and specificity will be discuss is shown in Table 3.

Accuracy represents the percentage of the correct classification for Poor, Average and Good classes. NBTree produces an accuracy of 94.1 % for Poor class. This means that approximately 94.1 % of the poor students are appearing as poor and NBTree classifier correctly reject those that are not poor. Sensitivity

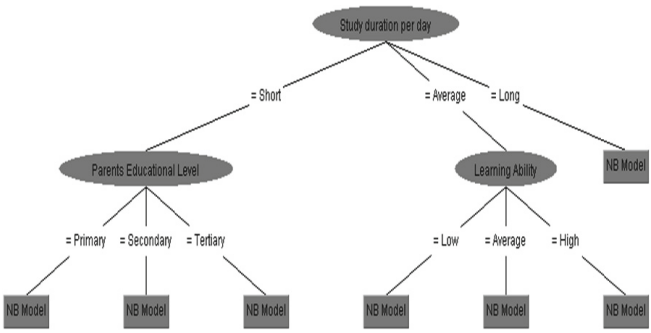


Fig. 3. Visualization of NBTree model

Table 3. Performance measure of NBTree classifier

Prediction	Accuracy (%)	Sensitivity (%)	Specificity (%)
Poor	94.1	80.0	97.3
Average	89.6	48.0	99.1
Good	88.1	98.8	70.0

measures how well the classifier predicts true when it is actually true. For example, all poor students are identified as poor and etc. Based on the result, NBTree produces a result of 80.0 % sensitivity or true positive rate for Poor class. This shows that NBTree classifier has a higher rate, which is 80.0 % to identify the students who are likely to fail.

Specificity measures how well the classifier predicts false when it is actually false. For example, no average or good students are identified as poor. Based on the result in Table 3, NBTree come out with a result of 97.3 % specificity or true negative rate of Poor class. This shows that NBTree classifier has a higher rate of 97.3 % to identify not poor student as average or good students.

4 Result and Discussion

The proposed hybrid algorithm which is NBTree is used to compare with the Decision Tree and Naïve Bayes classifiers when run using the identified parameters to determine the predictive accuracy of the algorithms. These classifiers is run using the 10-folds cross validation technique and the result of the three algorithms when run with identified parameters are shown in Table 4.

Table 4. Performance measure of NBTree classifier

Classifiers	Decision tree	Naïve Bayes	Specificity (%)
Prediction accuracy (%)	63.7	72.6	85.9

Based on the result, NBTree classifier produces a prediction accuracy of 85.9 %, where 116 out of 135 data instances are predicted accurately. For Naïve Bayes classifier, it gives a prediction accuracy of 72.6 % as 98 out of 135 data instances are correctly predicted. Lastly for Decision Tree classifiers, it gives a prediction accuracy of 63.7 % as it only correctly classified 86 instances.

NBTree algorithm shows a higher prediction accuracy compared to Naïve Bayes and Decision Tree algorithms because it utilizes the advantages of both Decision Tree (segmentation) and Naïve Bayes (evidence accumulation from multiple attributes) classifiers. It will segment the data using a univariate Decision-Tree and in each segment of the data, represented by a leaf described by the Naïve Bayes classifier. Decision Tree classifier is showing the lowest prediction

accuracy because of the slightly larger tree with 23 leaves and 43 nodes. This large size of the tree makes Decision Tree hard to be understood and thus having low prediction accuracy.

5 Conclusion and Future Work

Until today, the prediction of students' performance still remains as the major concern for any education sector. It is important as it helps the educators to identify and assist the students that are having the risk of poor performance. At the same time, it also helps the educators develop new teaching methods to adapt to students' different learning ability. This paper has identified a new sets of parameters for predicting the performance of new students. While for the prediction techniques, NBTree classifier has proven to have higher prediction accuracy which is 85.9 % compared to Decision Tree and Naïve Bayes classifiers which are having accuracy of 63.7 % and 72.6 % respectively. Even though this classifier prediction accuracy is only 85.9 %, but it is considered as a good classification method for educators to predict students' performance. The future works of this research are to try using others data mining methods and explore more parameters using different feature selection techniques and hopefully can improve the accuracy of current prediction method that gives only 85.9 %.

Acknowledgments. We would like to show our gratitude to Universiti Sains Malaysia (USM) for supporting this research.

References

1. Romero, C., Ventura, S.: Educational data mining-a review of the state of the art. IEEE Trans. Syst., Man, Cybern.- Part C: Appl. Rev. **40**(6), 601–619 (2010)
2. Farid, D.M., et al.: Hybrid decision tree and Naïve Bayes classifiers for multi-class classification tasks. Expert Syst. Appl. **41**(4), 1937–1946 (2014)
3. What is the difference between supervised learning, unsupervised learning? Stackoverflow.com (2016). http://stackoverflow.com/questions/1832076/what-is-the-difference-between-supervised-learning-and-unsupervised-learning
4. What is a Decision Tree? -Examples, Advantages & Role in Management - Video & Lesson Transcript | Study.com, Study.com (2015). http://study.com/academy/lesson/what-is-a-decision-tree-examples-advantages-role-in-management.html
5. Quadril, M.N., Kalyankar, N.V.: Drop out feature of student data for academic performance using decision tree techniques. Glob. J. Comput. Sci. Technol. **10**(2), 1–5 (2010)
6. Osmanbegovic, E., Suljic, M.: Data mining approach for predicting student performance. J. Econ. Bus. **10**(1), 51–56 (2012)
7. Naïve Bayes for Dummies a Simple Explanation, Text Analysis Blog | Aylien (2015). http://blog.aylien.com/post/120703930533/naive-bayes-for-dummies-a-simple-explanation
8. Dangi, A., Srivastava, S.: Educational data classification using selective Naive Bayes for quota categorization. In: IEEE International Conference on MOOC, Innovation, Technology in Education (MITE), pp. 118–121 (2014)

9. Shana, J., Venkatachalam, T.: Identifying key performance indicators and predicting the result from student data. Int. J. Comput. Appl. **25**(9), 45–48 (2011)
10. Qasem, A.A.-R., Emad, M.A.-S., Mustafa, I.A.-N.: Mining student data using decision trees. In: The 2006 International Arab Conference on Information Technology (ACIT 2006), pp. 1–5 (2006)
11. Kabra, R.R., Bichkar, R.S.: Performance prediction of engineering students using decision trees. Int. J. Comput. Appl. **36**(11), 8–12 (2011)
12. Mishra, T., Kumar, D., Gupta, S.: Mining students' data for prediction performance. In: Fourth International Conference on Advanced Computing & Communication Technologies, pp. 255–262 (2014)
13. Yadav, S.K., Pal, S.: Data mining- a prediciton of performance improvement using classification. World Comput. Sci. Inf. Technol. J. (WCSIT) **2**(2), 51–56 (2012)
14. Superby, J.F., Vandamme, J.P., Meskens, N.: Determination of factors influencing the achievement of the first-year university students using data mining methods. In: Workshop on Educational Data Mining, pp. 37–44 (2006)
15. Ibrahim, Z., Rusli, D.: Predicting students' academic performance-Comparing artificial neural network, decision tree, linear regression. In: 21st Annual SAS Malaysia Forum, pp. 1–6 (2007)
16. Ramaswami, M., Bhaskaran, R.: A CHAID based performance prediction model in educational data mining. Int. J. Comput. Sci. Issues, IJCSI **7**(1), 10–19 (2010)
17. Baradwaj, B.K., Pal, S.: Mining educational data to analyze students' performance. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **2**(6), 63–69 (2011)
18. Bresfelean, V.P., Bresfelean, M., Ghisoiu, N.: Determining students' academic failure profile founded on data mining methods. In: 30th International Conference on Information Technology Interfaces (ITI), pp. 317–322 (2008)
19. Shah, N.S.: Predicting factors that affect students' academic performance by using data mining techniques. Pak. Bus. Rev., 631–668 (2012)
20. Kovcic, Z.J.: Early prediction of student success- mining students' enrolment data. In: Proceedings of Informing Science & IT Education Conference (InSITE), pp. 647–665 (2010)
21. Cortez, P., Silva, A.: Using data mining to predict secondary school student performance. In: Brito, A., Teixeira, J. (eds.) Proceedings of 5th Future Business Technology Conference, pp. 5–12 (2008)
22. Ishitani, T.T.: A longitudinal approach to assessing attrition behavior among first-generation students: time-varying effects of pre-college characteristic. Res. High. Educ. **44**(4), 433–449 (2003)
23. Shahiri, A.M., Husain, W., Rashid, N.A.: A review on predicting student's performance using data mining techniques. In: The Third Information Systems International Conference, pp. 1–8 (2015)
24. Kabakchieva, D.: Predicting student performance by using data mining methods for classification. Cybern. Inf. Technol. **13**(1), 61–72 (2013)
25. Ramesh, V., Parkavi, P., Ramar, K.: Predicting student performance-a statistical and data mining approach. Int. J. Comput. Appl. **63**(8), 35–39 (2013)
26. Hijazi, S.T., Naqvi, S.M.M.R.: Factors affecting students' performance: a case of private colleges. Bangladesh e-J. Sociol. **3**(1), 90–100 (2006)
27. Khan, Z.N.: Scholastic achievement of higher secondary students in science stream. J. Soc. Sci. **1**(2), 84–87 (2005)
28. Romero, C., et al.: Data mining algorithms to classify students. In: Ist International Conference on Educational Data Mining Montreal, Quebec, Canada, pp. 8–17 (2008)

29. Jishan, S.T., et al.: Improving accuracy of students' final grade prediction model using optimal equal width binning and synthetic minority over-sampling technique. Dec. Analytics **2**(1), 1–25 (2015)
30. Shyamala, D., et al.: A study on feature selection, classification techniques for automatic genre classification of traditional malay music. In: ISMIR- Section 3A-Content Based Retrieval, Categorization, Similarity, pp. 331–336 (2011)
31. Shana, J., Venkatachalam, T.: Identifying key performance indicators and predicting the result from student data. Int. J. Comput. Appl. **25**(9), 45–48 (2011)
32. Dinakaran, S., Ranjit, J.T.P.: Role of attribute selection in classification algorithms. Int. J. Sci. Eng. Res. **4**(6), 67–71 (2013)
33. Kohavi, R.: Scaling up the accuracy of Naïve-Bayes classifiers: a decision-tree hybrid. In: Proceedings of the Second International Conference on Knowledge Discovery and Data Mining, pp. 202–207 (1996)
34. Jiang, L., Li, C.: Scaling up the accuracy of decision-tree classifiers: a Naïve-Bayes combination. J. Comput. **6**(7), 1325–1331 (2011)

An Improved Method for Stock Market Forecasting Combining High-Order Time-Variant Fuzzy Logical Relationship Groups and Particle Swam Optimization

Nghiem Van Tinh[1(✉)] and Nguyen Cong Dieu[2]

[1] Thai Nguyen University of Technology, Thai Nguyen University, Thai Nguyen, Vietnam
nghiemvantinh@tnut.edu.vn

[2] Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam
ncdieu@yahoo.com
http://www.tnut.edu.vn

Abstract. Fuzzy forecasting approaches are mainly based on the modeling of fuzzy logical relationships of the historical data. In this paper, an improved model for forecasting stock market indices which combines the High-order Time-Variant Fuzzy Logical Relationship Groups (HV-FLRGs) and Particle Swarm Optimization (PSO) is presented. Firstly, HV-FLRGs are more effective to capture fuzzy relations on time series data than the conventional time-invariant fuzzy logical relationship groups. Secondly, PSO is employed to optimize the length of intervals by searching the space of the universe of discourse. To verify the effectiveness of the proposed model, the historical data of Taiwan Futures Exchange (TAIFEX) are examined. The simulation result shows that the proposed model outperforms the previous forecasting models based on the high-order fuzzy time series. These results are very promising for the future work on the development of fuzzy time series and PSO algorithm in real-world forecasting applications.

Keywords: Forecasting · Fuzzy time series · Fuzzy relationship group · Particle swarm optimization · Stock market

1 Introduction

In order to advance the decision-making process concerning future requirements, many researchers have focused on real-world problems to deal with various time series data, such academic enrolments [3–5,7], temperature prediction [11], crop productions [10] and stock markets [12]. However, the conventional forecasting methods only refer to real numbers and fail to solve forecasting problems in which the historical data are represented by linguistic values. Fuzzy set theory was originally proposed by Zadeh [1] to handle human linguistic terms problems

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_18

and successfully applied to time series forecasting in [2]. They introduced the time-invariant fuzzy time series and the time-variant time series model which use the max–min operations to forecast the enrolments of the University of Alabama. Unfortunately, their method had many drawbacks such as huge computation when the fuzzy rule matrix is large and lack of persuasiveness in determining the universe of discourse and the length of intervals. Therefore, Ref. [5] proposed the first-order fuzzy time series model by using simple arithmetic calculations instead of max-min composition operations [2–4] for better forecasting accuracy. Thereafter, the fuzzy time series methods have received increasing attention in many forecasting applications. To achieve better forecasting accuracy, Ref. [6] presented an effective approach which can properly adjust the lengths of intervals. Chen in [7] presented a new forecasting model based on the high-order fuzzy logical relationship groups to forecast the enrolments of the University of Alabama. Yu in [8] proposed a new model which refined the lengths of intervals during the formulation of fuzzy relationships, thus capturing the fuzzy relations more effectively. Singh in [10] presented a simplified and robust computational method for the forecasting rules based on one and various parameters as fuzzy relationships. Lee et al. in [13] presented method for forecasting the temperature and the TAIFEX based on fuzzy logic relation groups and genetic algorithm. They also used genetic algorithm and simulated annealing. Recently, Particle swarm optimization (PSO) technique has been successfully applied in many applications. Huang et al. in [16] proposed a new forecasting model based on two computational methods, fuzzy time series and PSO, for academic enrolments. Based on Chen's model [5], Kuo et al. in [14] introduced a new hybrid forecasting model which combined fuzzy time series with PSO algorithm to find the proper length of each interval. Then, to improve method in [14], Kuo et al. in [15] presented a new hybrid forecast method to solve the TAIFEX forecasting problem based on fuzzy time series and PSO. Additionly, in [18] used a fuzzy relation matrix obtained also from PSO technique based on the high-order fuzzy time series. The above-mentioned researches showed that the lengths of intervals and fuzzy relations are two critical factors for forecasting accuracy. Therefore, we propose an improved forecasting method which combined the HV-FLRGs and PSO algorithm. Firstly, the proposed method fuzzifies the historical data into fuzzy sets to form high-order fuzzy logical relationships. Secondly, the PSO algorithm for the optimized lengths of intervals is developed to adjust interval lengths by searching the space of the universe of discourse. The case study with the data of TAIFEX shows that the performance of our model is better than those of any existing methods. The remainder of this paper is organized as follows. In Sect. 2, a brief review of the concepts of fuzzy time series and PSO algorithm are introduced. In Sect. 3, an improved forecasting model based on the HV-FLRGs and PSO is presented. Section 4, evaluates the forecasting performance of the proposed method with the existing methods on the historical data set of TAIFEX. Finally, some conclusions are discussed in Sect. 5.

2 Fuzzy Time Series and PSO Algorithm

2.1 Basic Concepts of Fuzzy Time Series

This section briefly summarizes the basic fuzzy and fuzzy time series concepts. The main difference between the fuzzy time series and traditional time series is that the values of the fuzzy time series are represented by fuzzy sets rather than real value. Let $U = \{u_1, u_2, \ldots, u_n\}$ be an universal set; a fuzzy set A of U is defined as $A = \{f_A(u_1)/u_1 + f_A(u_2)/u_2 + \ldots + f_A(u_n)/u_n\}$, where f_A is a membership function of a given set A: $U \rightarrow [0, 1]$, $f_A(u_i)$ indicates the grade of membership of u_i in the fuzzy set A. $f_A(u_i) \in [0, 1]$, and $1 \leq i \leq n$. General definitions of fuzzy time series are given as follows:

Definition 1. Fuzzy time series Let $Y(t)(t = ..,0,1,2..)$, a subset of R, be the universe of discourse on which fuzzy sets $f_i(t)(i = 1, 2, \ldots)$ are defined and if F(t) be a collection of $f_1(t), f_2(t), \ldots$, then F(t) is called a fuzzy time series on $Y(t)(t \ldots, 0, 1, 2, \ldots)$.

Definition 2. Fuzzy logic relationships (FLRs) If there exists a fuzzy relationship R(t-1,t), such that $F(t) = F(t-1) * R(t-1, t)$, where "$*$" is an arithmetic operator, then F(t) is said to be caused by F(t-1). The relationship between F(t) and F(t-1) can be denoted by $F(t-1) \rightarrow F(t)$. Let $A_i = F(t)$ and $A_j = F(t-1)$, the relationship between F(t) and F(t-1) is denoted by fuzzy logical relationship $A_i \rightarrow A_j$ where A_i and A_j refer to the current state or the left - hand side and the next state or the right-hand side of fuzzy time series.

Definition 3. λ- Order Fuzzy Relations
Let $F(t)$ be a fuzzy time series. If $F(t)$ is caused by $F(t-1), F(t-2), \ldots, F(t-\lambda+1)F(t-\lambda)$ then this fuzzy relationship is represented by $F(t-\lambda), \ldots, F(t-2), F(t-1) \rightarrow F(t)$ and is called an λ- order fuzzy time series.

Definition 4. Fuzzy Relationship Group (FLRG)
Fuzzy logical relationships, which have the same left-hand sides, can be grouped together into fuzzy logical relationship groups. Suppose there are relationships such that

$$A_i \rightarrow A_k$$
$$A_i \rightarrow A_m$$
$$\ldots.$$

So, based on [5], these fuzzy logical relationship can be grouped into the same FLRG as : $A_i \rightarrow A_k, A_m \ldots$

Definition 5. Time-Variant Fuzzy Relationship Group (TV-FLRG)
The fuzzy relationship is determined by the relationship of $F(t-1)ßF(t)$. If, let $F(t) = A_i(t)$ and $F(t-1) = A_j(t-1)$, we will have the relationship $A_j(t-1) \rightarrow A_i(t)$. At the time t, we have the following fuzzy relationships: $A_j(t-1) \rightarrow A_i(t), A_j(t1-1) \rightarrow A_{i_1}(t1), \ldots, A_j(tp-1) \rightarrow A_{i_p}(tp)$

with $t1, t2, \ldots, tp \leq$ t. It means that if the fuzzy relationship took place before $A_j(t-1) \rightarrow A_i(t)$, we can group the fuzzy logic relationship to be $A_j(t-1) \rightarrow A_{i_1}(t1), A_{i_2}(t2), A_{i_p}(tp), A_i(t)$. It is called time-variant fuzzy logic relationship group

2.2 PSO Algorithm

PSO was first introduced by Eberhart and Kannedy in 1995. It belongs to a population-based evolutionary algorithm that can efficiently search a nearly optimal or optimal solution for optimization problems. Most population-based approaches are motivated by evolution as seen in nature. The development of PSO algorithm [14–16] was inspired by the social behaviour of animals, such as fish schooling, birds flocking and the swarm theory. The PSO algorithm applies a cooperative particle swarm to find the best solution from all feasible solutions. Each particle is randomly initialized and then allowed to move in the virtual searching space. At each step of optimization, each particle evaluates its own fitness and the fitness of its neighbouring particles. Each particle can remember its own best solution, which results in the best fitness, as well as see the candidate solution for the best performing particle in its neighbourhood. A moving particle, indexed by id, adjusts its candidate solution according to the following formulas:

$$V_{id}^{k+1} = \omega^k * V_{id}^k + C_1 * Rand() * (Pbest_{id} - X_{id}^k) + C_2 * Rand() * (G_{best} - X_{id}^k) \quad (1)$$

$$X_{id}^{k+1} = X_{id}^k + V_{id}^{k+1} \quad (2)$$

$$\omega^k = \omega_{max} - k * (\omega_{max} - \omega_{min})/(iter_{max}) \quad (3)$$

where, X_{id}^k is the current position of a particle id in $^k th$ iteration; V_{id}^k is the velocity of the particle id in k^{th} iteration, and is limited to $[-V_{max}, V_{max}]$ where V_{max} is a constant pre-defined by user. *Pbest* is the position of the particle that experiences the best fitness value. G_{best} is the best one of all personal best positions of all particles within the swarm. Rand() is the function can generate a random real number between 0 and 1 under normal distribution. C_1 and C_2 are acceleration values which represent the selfcondence coefficient and the social coefficient, respectively.
The steps for the standard PSO are presented in Algorithm 1.

Algorithm 1. Standard PSO algorithm
Step 1. initialize all particles' positions X_{id} and velocities V_{id}
Step 2. **while** the stop condition (the maximal moving steps are reached) is not satisfied **do**
2.1 .**for** particle i, ($1 \leq i \leq$ NumberOfParticles) **do**
calculate the fitness value of particle i
update the personal best position of particle i according to the fitness value
end **for**
2.2. update the global best position of all particles according to the fitness value.

2.3. **for** particle i, $(1 \leq i \leq NumberOfParticles)$ **do**
move particle i to another position according to (1) and (2)
end **for**
end **while**

3 An Improved Forecasting Model Based on the HV-FLRGs and PSO Algorithm

Based on Kuo et al. in [14], a new forecasting model which combined the HV-FLRGs and PSO algorithm is introduced. In the proposed model, three key aspects have been applied to approach the lengths of intervals and fuzzy relations on time series data to increase forecasting accuracy. First, original historical data are used instead of the variations of historical data in our forecasting model. Second, the HV-FLRGs are derived from the time-variant fuzzy relationship groups and calculate the forecasting output based on the fuzzy sets on the right-hand side of the HV-FLRGs. Third, the PSO algorithm is developed to adjust the interval lengths to obtain the optimal partition. A detailed explanation of the proposed model in Subsect. 3.1 follow.

3.1 Forecasting Model Based on HV-FLRGs

To verify the effectiveness of the proposed model, the empirical data for the TAIFEX [13] (Historical data of the TAIFEX under 8/3/1998 – 8/31/1998) are used to illustrate the high - order fuzzy time series forecasting process. The step-wise procedure of the proposed model is detailed as follows:

Step 1: Define the universe of discourse U Assume Y(t) be the historical data of TAIFEX on date t. The university of discourse is defined as $U = [\beta_{min} - N_1, \beta_{max} + N_2]$ where β_{min}, β_{max} are The minimum and maximum data of $Y(t)$; N_1, and N_2 are two proper positive integers to tune the lower bound and upper bound of the U. From the historical data shown in [13], we obtain $\beta_{min} = 6566$ and $\beta_{max} = 7560$. Thus, the universe of discourse is defined as U= $[6500, 7600]$ with $N_1 = 66$ and $N_2 = 40$.

Step 2: Partition U into appropriate intervals
Divide U into equal length intervals. Compared to the previous models in [5,8], we cut U into seven intervals, $u_1, u_2, \ldots, u_7$, respectively. The length of each interval is $l = (\beta_{max} + N_2 - \beta_{min} + N_1) \div 7 = (7600 - 6500) \div 7 \approx 157$. Thus, the seven intervals are: $u_1 = (6500, 6657], u_2 = (6657, 6814], \ldots, u_6 = (7285, 7442], u_7 = (7442, 7600]$.

Step 3: Define the fuzzy sets.
Each interval in Step 2 represents a linguistic variable of "stock market" in [15]. For seven intervals, there are seven linguistic values which are A_1 = "*worst*", A_2 = "*bad*", A_3 = "*alittlebad*", A_4 = "*average*", A_5 = "*good*", A_6 = "*very good*", and A_7 = "*excellent*" to represent different regions in the

universe of discourse on U, respectively. Each linguistic variable represents a fuzzy set $A_i (1 \leq i \leq 7)$ and its definition is described in (4).

$$A_i = \sum_{i=1}^{7} \frac{a_{ij}}{u_j} = \begin{cases} 1 & \text{if } i == j \\ 0.5 & \text{if } j == i-1 \text{ or } j = i+1 \\ 0 & \text{otherwise} \end{cases} \tag{4}$$

where $a_{ij} \in [0,1], 1 \leq i \leq 7, 1 \leq j \leq 7$ and u_j is the ^{j}th interval of u. The value of a_{ij} indicates the grade of membership of u_j in the fuzzy set A_i.

Step 4: Fuzzy all historical data In order to fuzzify all historical data, it's necessary to assign a corresponding linguistic value to each interval first. The simplest way is to assign the linguistic value with respect to the corresponding fuzzy set that each interval belongs to with the highest membership degree. For example, the historical data on date 8/3/1998 is 7552, and it belongs to interval u_7 because 7552 is within (7442, 7600]. So, we then assign the linguistic value "excellent" (eg. the fuzzy set A_7) corresponding to interval u_7 to it. Consider two time serials data $Y(t)$ and $F(t)$ on date t, where $Y(t)$ is actual data and $F(t)$ is the fuzzy set of Y(t). According to formula (4), the fuzzy set A_7 has the maximum membership value at the interval u_7. Therefore, the historical data time series on date $Y(8/3/1998)$ is fuzzified to A_7. The completed fuzzified results of the TAIFEX are listed in Table 1.

Step 5: Create all λ-order fuzzy relationships
Based on Definition 3. To establish a λ-order fuzzy relationship, we should find out any relationship which has the $F(t-\lambda), F(t-\lambda+1), ..., F(t-1)F(t)$, where $F(t-\lambda), F(t-\lambda+1), \ldots, F(t-1)$ and F(t) are called the current state and the next state, respectively. Then, a λ - order fuzzy relationship is got by replacing the corresponding linguistic values. For example, supposed $\lambda = 3$, a fuzzy relationship $A_7, A_7, A_7 \rightarrow A_7$ is got as $F(8/3/1998), F(8/4/1998)F(8/5/1998) \rightarrow F(8/6/1998)$. So, from Table 1. we get ^{3}rd-order fuzzy relationships are shown in Table 2.

Step 6: Establish all λ-order fuzzy relationships groups
By [5,14], all the fuzzy relationship having the same fuzzy set on the left-hand side or the same current state can be put together into one fuzzy relationship group. But, according to the Definition 5, we need to consider the appearance history of the fuzzy sets on the right-hand side too. Therefore, only the element on the right hand side appearing before the left-hand side of the relationship group is taken into the same fuzzy logic relationship groups. Thus, from Table 2 and based on Definition 5, we can obtain 20 fuzzy logical relationship groups shown in Table 3.

Step 7: Calculate the forecasting output
In this step, we create all forecast outputs for fuzzy logical relationship groups based on fuzzy sets on the right-hand or next state within the same group. For each group in Table 3, we divide each corresponding interval of each next state

Table 1. The results of fuzzification

Year	Actual data	Fuzzy sets	Membership
8/3/1998	7552	A7	[0 0 0 0 0 0.5 1]
8/4/1998	7560	A7	[0 0 0 0 0 0.5 1]
8/5/1998	7487	A7	[0 0 0 0 0 0.5 1]
8/6/1998	7462	A7	[0 0 0 0 0 0.5 1]
8/7/1998	7515	A7	[0 0 0 0 0 0.5 1]
8/10/1998	7365	A6	[0 0 0 0 0.5 1 0.5]
8/11/1998	7360	A6	[0 0 0 0 0 .5 1 0.5]
8/12/1998	7330	A6	[0 0 0 0 0.5 1 0.5]
8/13/1998	7291	A6	[0 0 0 0 0 .5 1 0.5]
8/14/1998	7320	A6	[0 0 0 0 0.5 1 0.5]
8/15/1998	7300	A6	[0 0 0 0 0.5 1 0.5]
8/17/1998	7219	A5	[0 0 0 0.5 1 0.5 0]
8/18/1998	7220	A5	[0 0 0 0.5 1 0.5 0]
8/19/1998	7285	A5	[0 0 0 0.5 1 0.5 0]
8/20/1998	7274	A5	[0 0 0 0.5 1 0.5 0]
8/21/1998	7225	A5	[0 0 0 0.5 1 0.5 0]
8/24/1998	6965	A3	[0 0.5 1 0.5 0 0 0]
8/25/1998	6949	A3	[0 0.5 1 0.5 0 0 0]
8/26/1998	6790	A2	[0.5 1 0.5 0 0 0 0]
8/27/1998	6835	A3	[0 0.5 1 0.5 0 0 0]
8/28/1998	6695	A2	[0.5 1 0.5 0 0 0 0]
8/29/1998	6728	A2	[0.5 1 0.5 0 0 0 0]
8/31/1998	6566	A1	[1 0.5 0 0 0 0 0]

Table 2. Fuzzy logical relationships

Number	The ions	Number	The lations
1	A7, A7, A7 → A7	11	A6, A5, A5 → A5
2	A7, A7, A7 → A7	12	A5, A5, A5 → A5
3	A7, A7, A7 → A6	13	A5, A5, A5 → A5
4	A7, A7, A6 → A6	14	A5, A5, A5 → A3
5	A7, A6, A6 → A6	15	A5, A5, A3 → A3
6	A6, A6, A6 → A6	16	A5, A3, A3 → A2
7	A6, A6, A6 → A6	17	A3, A3, A2 → A3
8	A6, A6, A6 → A6	18	A3, A2, A3 → A2
9	A6, A6, A6 → A5	19	A2, A3, A2 → A2
10	A6, A6, A5 → A5	20	A3, A2, A2 → A1

Table 3. ^3rd-oder Fuzzy logical relationship groups

No group	^3rd-oder Fuzzy relation groups	No group	^3rd-oder Fuzzy relation groups
G1	A7	G11	A5
G2	A7, A7	G12	A5
G3	A7, A7, A6	G13	A5, A5
G4	A6	G14	A5, A5, A3
G5	A6	G15	A3
G6	A6	G16	A2
G7	A6, A6	G17	A3
G8	A6, A6, A6	G18	A2
G9	A6, A6, A6, A5	G19	A2
G10	A5	G20	A1

into p sub-regions with equal size, and create a forecasted value for each group according to equal (5).

$$forecasted = \frac{1}{n}\sum_{j=1}^{n}\frac{(m_{kj} + subm_{kj})}{2} \tag{5}$$

where;

- n is the total number of next states or the total number of fuzzy sets on the right-hand side within the same group

- m_{kj} $(1 \leq j \leq n)$ is the midpoint of interval u_{kj} corresponding to ^jth fuzzy set on the right-hand side where the highest level of fuzzy set A_{kj} takes place in these intervals, u_{kj}.

- $subm_{kj}$ is the midpoint of one of p sub-regions corresponding to ^jth fuzzy set on the right-hand side where the highest level of A_{kj} takes place in this interval. Based on equal (5) and the data in Table 1, we obtain forecasted results for TAIFEX from 8/3/1998 to 9/1/1998 based on ^3rd-order fuzzy time series model with seven intervals are listed in Table 4.

To calculate the forecasted performance of proposed method in the fuzzy time series, the mean square error (MSE) are used as an evaluation criterion to represent the forecasted accuracy. The MSE value is computed according to (6) as follows

$$MSE = \frac{1}{n}\sum_{i=\lambda}^{n}(Fo_i - Ac_i)^2 \tag{6}$$

Where, Ac_i denotes actual data on date i, Fo_i is forecasted value on date i, n is number of the forecasted data, λ is order of the fuzzy relationships

Table 4. The complete forecasted outputs based on ^3rd-oder fuzzy time series model

Date	Actual data	Fuzzy set	Forecasted value
8/3/1998	7552	A7	Not forecasted
8/4/1998	7560	A7	Not forecasted
8/5/1998	7487	A7	Not forecasted
8/6/1998	7462	A7	7495.2
8/7/1998	7515	A7	7508.2
8/10/1998	7365	A6	7460.2
———	——	—	——
———	——	—	——
8/28/1998	6695	A2	6709.4
8/29/1998	6728	A2	6735.9
8/31/1998	6566	A1	6578.8

3.2 Forecating Method Based on the HV-FLRGs and PSO

To improve forecasted accuracy of the proposed, the effective lengths of intervals and time-variant fuzzy relationship groups which are two main issues presented in this paper. A novel method for forecasting TAIFEX is developed to adjust the length each of intervals in the universe of discourse without increasing the number of intervals by minimizing the MSE value. In our model, each particle exploits the intervals in the universe of discourse of historical data $Y(t)$. Let the number of the intervals be n, the lower bound and the upper bound of the universe of discourse on historical data $Y(t)$ be p_0 and p_n, respectively. Each particle is a vector consisting of n-1 elements p_i where $1 \leq i \leq n-1$ and $p_i \leq p_{(i+1)}$. Based on these n-1 elements, define the n intervals as $u_1 = [p_0, p_1], u_2 = [p_1, p_2], \cdots, u_i = \left[p_{(i-1)}, p_i\right], \cdots, u_n = \left[p_{(n-1)}, p_n\right]$ respectively. When a particle moves to a new position, the elements of the corresponding new vector need to be sorted to ensure that each element $p_i (1 \leq i \leq n-1)$ arranges in an ascending order. The step-wise procedure of the proposed method is illustrated in Algorithm 2.

Algorithm 2. The HV-FLRGs-PSO algorithm

1. initialize all particles' positions X_{id} and velocities V_{id}
2. **while** the stop condition (maximum iterations or minimum MSE criteria) is not satisfied **do**

2.1. **for** particle i, $(1 \leq i \leq NumberOfParticles)$ **do**

- partition U into new intervals by Step 2 in 3.1
- define linguistic terms by step 3 in 3.1
- fuzzify all historical data by Step 4 in 3.1
- create all λ - order fuzzy relationships by Step 5
- Establish all λ - order fuzzy relationship groups by Step 6
- calculate forecasting values by Step 7

• compute the MSE values for particle i based on (6)
• update the personal best position of particle i according to the MSE values
end **for**
2.2. update the global best position of all particles according to the MSE values
3. **for** particle i, $(1 \leq i \leq NumberOfParticles)$ **do**
move particle i to another position according to (1) and (2)
end **for**
update ω according to (3)
end **while**

4 Computational Results

4.1 Preliminary Data

In this paper, we apply the proposed method to forecast TAIFEX index with the whole historical data [13], from 8/3/1998 to 9/30/1998 are used to perform comparative study in the training phase. The essential parameters of proposed model for forecasting TAIFEX are listed in Table 5.

4.2 Computational Results

In order to verify the forecasting effectiveness of the proposed model with the high-order FLRGs and different numbers of intervals, six FTS models in C96 [5], H01b [6], L06 [9], L08 [13], HPSO [14] and the NPSO [15], are examined and compared. The forecasted accuracy of the proposed method is estimated using the MSE technique in (6). The simulation result is expressed in Table 5. Our proposed model is executed 10 runs, and the best result of runs is taken to be the final result. A comparison of the fitness accuracy (i.e. the MSE value) with various orders and different number of intervals among the proposed model, the C96 model, H01b model, the L06 model, the L08 model, the HPSO model and the NPSO model are listed in Table 6.

From Table 6, it is obvious that our model has a smaller MSE value than the other fuzzy forecasting models. The MSE value is calculated according to equal

Table 5. Parameters used for forecasting TAIFEX

Number of particles	30
Maximum number of iterations	100
The value of ω be linearly decreased	0.9 to 0.4
The coefficient C1 = C2	2
The velocity be limited to	[−100,100]
The position be limited to	[6200,7600]

Table 6. A comparison of the forecasted results of the proposed method with the existing models based on high – order fuzzy time series under number of intervals = 16

Date	Actual data	C96	H01b	L06	L08	HPS0	NPSO	Our method
8/3/1998	7552	-	-	-	-	-	-	-
8/4/1998	7560	7450	7450	-	-	-	-	-
8/5/1998	7487	7450	7450	-	-	-		-
8/6/1998	7462	7500	7500	7450	-	-	7452.54	-
8/7/1998	7515	7500	7500	7550	-	-	7331.62	-
8/10/1998	7365	7450	7450	7350	-	-	7285.63	-
8/11/1998	7360	7300	7300	7350	-	-	7331.62	-
8/12/1998	7330	7300	7300	7350	7329	7289.56	7291.67	7329.78
8/13/1998	7291	7300	7300	7250	7289.5	7320.77	7217.15	7287.22
8/14/1998	7320	7183.33	7188.33	7350	7329	7289.56	7217.15	7324.78
———	———	———	———	———	———	———	———	———
9/28/1998	6840	6850	6750	6850	6848	6800.07	7452.54	6839.19
9/29/1998	6806	6850	6850	6850	6796	6800.07	7331.62	6799.74
9/30/1998	6787	6850	6750	6750	6796	7289.56	7285.63	6785.24
MSE		9668.94	5437.58	1364.56	105.02	103.61	35.86	**18.49**

(6) following:

$$MSE = \frac{1}{40}\sum_{i=7}^{40}(7329.78 - 7330)^2 + \cdots + (6785.24 - 6787)^2 = 18.49$$

In addition, we also perform five more runs with different orders and 16 intervals of the universe set of discourse to be compared with other models such as L08 in [13] (based on GA), HPSO in [14] (based on PSO), and NPSO [15] (also based on PSO). The detail of comparison is shown in Table 7. The forecasting trend is depicted in Fig. 1 for clearer illustration.

During the simulation, the number of intervals is kept for the existing models and our model. A comparing of MSE value is listed in Table 7. In Table 7, it can be seen that the accuracy of the proposed model is improved significantly. Particularly, our model gets the lowest MSE value of 18.49 with 7th-order fuzzy relation and the average MSE value of the proposed model is 24.22, which is smallest among four forecasting models. In addition, we also rebuilt NPSO model [14] is considered to be quite effective in recent years and compare the forecasting accuracy of this model with the proposed model on the same historical data of the TAIFEX with different number of samples as 15, 20, 25, 30, 35, 40, 45 and 47. The detail is presented in Table 8

From Table 8, it can be seen that our proposed model gives remarkably better forecasting accuracy with MSE values compared to NPSO model with the different number of samples

Table 7. A comparison of the forecasted results of the proposed method with the existing models based on high – order fuzzy time series under number of intervals = 16

Order	L08 model	HPSO model	NPSO model	Our model
3	208.79	152.47	35.86	28.58
4	142.26	148.14	34.76	27.12
5	143.61	112.24	29.30	25.72
6	147.14	122.68	23.95	21.31
7	105.02	103.61	25.56	**18.49**
8	124.48	108.37	24.7	24.13
Average	145.17	124.58	29.02	**24.22**

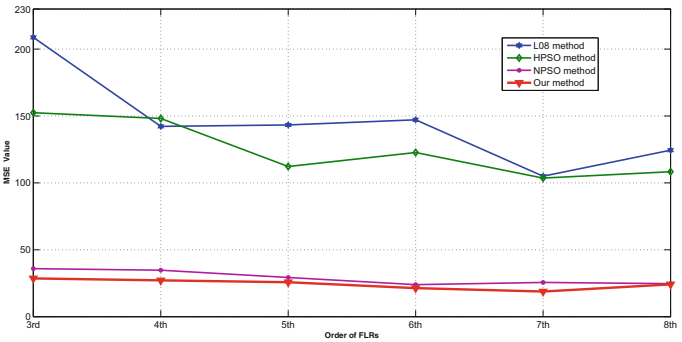


Fig. 1. A comparison of the MSE value between our model and the previous methods: L08, HPSO, NPSO based on high –order FTS with number of intervals =16.

Table 8. The comparison(MSE) between the proposed model and NPSO model on the same historical data set but different in the numbers of sampling values

Models	Number of historical data							
	15	20	25	30	35	40	45	47
NPSO model	0.47	3.75	9.01	13.4	18.6	25.47	32.38	36
Our method	0.38	2.24	8.4	13.4	17.29	21.86	29.15	28.58

5 Conclusion and Discussion

Stock market indices are very volatile time series in nature and it has difficult to make the potential relationship as a mathematical model. So, fuzzy time series has shown good performances for these real world problems. In order to improve the forecasting accuracy of the NPSO model, we consider the appearance history of the fuzzy sets on the right-hand side of the same fuzzy relation to create time–variant fuzzy logic relationship groups. Also we consider more information within

all next states of all fuzzy relationships to calculate the forecasting output for these fuzzy relationship groups. Then, a novel hybrid forecasting model based on an aggregated HV-FLRGs and PSO is developed to adjust the length of each interval in the universe of discourse. After applying the proposed forecasting method for the real world datasets of TAIFEX, we found that our approach shows better forecasting accuracy than previous ones. The detail of comparison was presented in Tables 6, 7 and 8. The main contributions of this paper are illustrated in the following. First, we show the forecasted accuracy is affected by calculating the forecasting rules from time-variant fuzzy relationship groups. Second, the computational results show that the proposed model gets highest forecasted accuracy for with seven - order FTS model. Actually, as listed in Table 7, the minimal MSE value for the proposed model is 18.49 which is the lowest forecasting accuracy among the models as shown in Table 6. Finally, our forecasting method is general enough for different kinds of time series and can be used in various applications efficiently.

Acknowledgments. The work reported in this paper has been supported by Thai Nguyen University of Technology - Thai Nguyen University. We also acknowledge the anonymous reviewers for comments that lead to clarification of the paper.

References

1. Zadeh, A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
2. Song, Q., Chissom, B.S.: Fuzzy time series and its models. Fuzzy Sets Syst. **54**(3), 269–277 (1993)
3. Song, Q., Chissom, B.S.: Forecasting enrolments with fuzzy time series - part I. Fuzzy Sets Syst. **54**(1), 1–9 (1993)
4. Song, Q., Chissom, B.S.: Forecasting enrolments with fuzzy time series - part II. Fuzzy Sets Syst. (1994)
5. Chen, S.M.: Forecasting enrolments based on fuzzy time series. Fuzzy Sets Syst. **81**, 311–319 (1996)
6. Huarng, K.: Effective lengths of intervals to improve forecasting in fuzzy time series. Fuzzy Sets Syst. **123**(3), 387–394 (2001)
7. Chen, S.M.: Forecasting enrolments based on high-order fuzzy time series. Cybern. Syst. **33**(1), 1–16 (2002)
8. Yu, H.K.: A refined fuzzy time-series model for forecasting. Phys. A: Stat. Mech. Appl. **346**(3–4), 657–681 (2005)
9. Lee, L.W., Wang, L.H., Chen, S.M., Leu, Y.H.: Handling forecasting problems based on two-factors high-order fuzzy time series. IEEE Trans. Fuzzy Syst. **14**, 468–477 (2006)
10. Singh, S.R.: A simple method of forecasting based on fuzzy time series. Appl. Math. Comput. **186**, 330–339 (2007a)
11. Lee, L.-W., Wang, L.-H., Chen, S.-M.: Temperature prediction and TAIFEX forecasting based on fuzzy logical relationships and genetic algorithms. Expert Syst. Appl. **33**, 539–550 (2007b)
12. Chen, T.-L., Cheng, C.-H., Teoh, H.-J.: High-order fuzzy time-series based on multi-period adaptation model for forecasting stock markets. Phys. A: Stat. Mech. Appl. **387**, 876–888 (2008a)

13. Lee, L.-W., Wang, L.-H., Chen, S.-M.: Temperature prediction and TAIFEX forecasting based on high order fuzzy logical relationship and genetic simulated annealing techniques. Expert Syst. Appl. **34**, 328–336 (2008b)
14. Kuo, I.H., et al.: An improved method for forecasting enrolments based on fuzzy time series and particle swarm optimization. Expert Syst. Appl. **36**, 6108–6117 (2009)
15. Kuo, I.-H., Horng, S.-J., Chen, Y.-H., Run, R.-S., Kao, T.-W., Chen, R.-J., Lai, J.-L., Lin, T.-L.: Forecasting TAIFEX based on fuzzy time series and particle swarm optimization. Expert Syst. Appl. **2**(37), 1494–1502 (2010)
16. Huang, Y.-L., Horng, S.-J., He, M., Fan, P., Kao, T.-W., Khan, M.K., Lai, J.-L., Kuo, I.-H.: A hybrid forecasting model for enrollments based on aggregated fuzzy time series and particle swarm optimization. Expert Syst. Appl. **7**(38), 8014–8023 (2011)
17. Wei, L.-Y.: A GA-weighted ANFIS model based on multiples stock market volatility causality for TAIEX forecasting. Appl. Soft Comput. **13**, 911–920 (2013)
18. Egrioglu, E.: PSO-based high order time invariant fuzzy time series method. Econ. Model. **38**, 633–639 (2014)

An Iterative Method to Solve Boundary Value Problems with Irregular Boundary Conditions

Vu Vinh Quang(✉) and Truong Ha Hai

College of Information and Communication Technology,
Thai Nguyen University, Thai Nguyen, Vietnam
{vvquang,thhai}@ictu.edu.vn

Abstract. In this paper, we investigate a general model of boundary value problems with irregular boundary conditions, a model attracting a lot of attention from researchers all over the world [2–6]. By the two approaches of finding values of functions or derivatives on the unidentified boundary conditions, we propose two iterative schemes to define approximation solutions of the problem and compare the rate of convergence of the two iterative schemes. The performed numerical experiments show the effectiveness of the two proposed methods.

Keywords: Boundary value problems · Irregular boundary conditions · Iterative schemes · Elliptic equations

1 Introduction

We consider the heat conduction problem proposed by the author in [3]. Assume that temperature field, $u(x)$, satisfies the equation

$$(\nabla^2 - k^2)u(x) = 0; x = (x_1, x_2, ..., x_d) \in \Omega, \tag{1}$$

where $\Omega \subset R^d$ is open bounded Lipschitz domain, d the dimension of the space. Assume that Ω is bounded by a curve $\partial\Omega$ so that $\partial\Omega = \Gamma_1 \cup \Gamma_2; \Gamma_1, \Gamma_2 \neq \emptyset; \Gamma_1 \cap \Gamma_2 = \emptyset$.

The above differential equation describes the heat conduction model of a fin where $u(x)$ is the dimensional local fin temperature, $k^2 = h/\left(\widetilde{k}\delta_f\right)$, h is the surface heat transfer coefficient $\left[W/(m^2K)\right]$, $\tilde{k}$ is the thermal conductivity of the fin and δ_f is the half-fin thickness $[m]$. Let $n(x) = (n_1(x), n_2(x), ...n_d(x))^T$ be the the outward normal vector at $x \in \partial\Omega$ and $q(x) = \nabla u(x).n(x)$ be the normal heat flux at a point $x \in \partial\Omega$. In formulating the problem, description of constant k, the location, shape and size of the entire boundary $\partial\Omega$, temperature and normal heat flux on the entire boundary $\partial\Omega$ are represented by Dirichlet, Neumann boundary conditions or mixed boundary conditions to enable the determination of heat distribution in the entire domain.

In reality, a special situation arises when it is possible to measure both the temperature and the normal heat flux on a part of the boundary $\partial\Omega$, say Γ_1,

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_19

at the same time it is impossible to measure the temperature and the normal heat flux on the rest part of the boundary, Γ_2. This leads to the mathematical formulation of Cauchy problem described by the partial differential equation and the boundary conditions

$$u(x) = \tilde{u}(x), q(x) = \tilde{q}(x), x \in \Gamma_1, \tag{2}$$

where $\tilde{u} \in H^{1/2}(\Gamma_1)$ and $\tilde{q} \in \left(H^{1/2}(\Gamma_1)\right)^*$ are prescribed temperature and normal heat flux, respectively. Because of special boundary conditions (Γ_1 is over-specified and Γ_2 is under-specified), the above problem is often called boundary value problem with incomplete boundary conditions or irregular boundary conditions. We assume that data are chosen such that there exists a solution to Cauchy problem and this solution is unique. A necessary condition for the Cauchy problem given by Eqs. (1)–(2) to be identifiable is that $\text{meas}(\Gamma_1) \geq \text{meas}(\Gamma_2)$.

Although this problem may have a unique solution, this solution is unstable with respect to small perturbations into the data into because of experimental instruments. Therefore, this problem is ill-posed and we cannot use direct approaches such as the least-square method in order to solve the system of linear equations which arises from the discretisation of the partial differential equations and the boundary conditions by multi-grid methods and finite element methods. Therefore, it is necessary to study methods to precisely solve the problems (1)–(2).

2 Alternating Iterative Algorithm

In the theory of partial differential equations, the boundary value problem can be solved if it is possible to determine values of boundary conditions on the entire boundary. Boundary conditions can be in the form of a function (Dirichlet), a normal derivative (Neumann) or mixed boundary conditions of function and normal derivative (Robin). Therefore, to be able to solve the discussed ill-posed problem, we need to determine values of boundary conditions on the rest of the boundary Γ_2, which is unknown during the experimental process with instruments (conditions can be Dirichlet or Neumann). One of the approaches is that through the existing conditions on Γ_1, we can determine normal derivative boundary conditions on Γ_2 by iterative algorithm. An alternating iterative algorithm for the simultaneous reconstruction of the unknown temperature $u|_{\Gamma_2}$ and normal heat flux $q|_{\Gamma_2}$ on the under-specified boundary proposed by Kozlov et al. [2]. Alternating iterative algorithm allows us to solve the two well-posed boundary value problems with mixed boundary conditions. The solutions of these problems can be easily determined by systems of sample programs in the library. Similarly we can develop another alternating iterative algorithm on the basis of determining the value of derivative on Γ_2.

In [4], Marin L. proposed MFS method to determine approximation solutions of the problem by using system of basic functions and the problem solution is found by solving system of linear algebraic equations in order to determine expansion coefficients. However, it can be seen that MFS method can be only applied when the right-hand side of the differential Eq. (1) is homogeneous. In

the case that it is not, we cannot determine the corresponding system of basic functions. Based on the method of determining iterative schemes boundary operator in combination with domain decomposition method, in this paper we will present results of the development of iterative schemes to determine approximate solutions of the problem in the general case when the right-hand side is different from 0.

3 Proposed Iterative Schemes

We consider the general problem

$$\begin{aligned} \Delta u - c^2 u &= f(x), \quad x \in \Omega, \\ \frac{\partial u}{\partial n} &= g(x), \quad x \in \Gamma_1, \\ u &= \varphi(x), \quad x \in \Gamma_1. \end{aligned} \tag{3}$$

This is the model of the boundary value problem with irregular boundary condition when the right-hand side $f(x) \neq 0$. Knowing that MFS method can be only applied when $f(x) \equiv 0$. To solve the this problem, we need to determine values of conditions on Γ_2 to transform the problem (3) into elliptic boundary value problem with weak or strong mixed boundary conditions. In order to determine values of conditions on Γ_2, there are two approaches:

- The first approach: determine $u|_{\Gamma_2}$
- The second approach: determine $\dfrac{\partial u}{\partial n}|_{\Gamma_2}$

3.1 Approach to Determine $u|_{\Gamma_2}$

Let $\xi = u|_{\Gamma_2}$, we consider the following iterative schemes for finding x:
Iterative Schemes QH1: With $\xi^{(0)} = 0, \forall k = 0, 1, 2, ...$ we solve the two problems respectively

- Step 1: Solve the problem with $u^{(k)}$

$$\begin{aligned} \Delta u^{(k)} - c^2 u^{(k)} &= f(x), \quad x \in \Omega, \\ \frac{\partial u^{(k)}}{\partial n} &= g, \quad x \in \Gamma_1, \\ u^{(k)} &= \xi^{(k)}, \quad x \in \Gamma_2. \end{aligned} \tag{4}$$

- Step 2: Solve the problem with $v^{(k)}$

$$\begin{aligned} \Delta v^{(k)} - c^2 v^{(k)} &= f(x), \quad x \in \Omega, \\ \frac{\partial v^{(k)}}{\partial n} &= \frac{\partial u^{(k)}}{\partial n}, \quad x \in \Gamma_2, \\ v^{(k)} &= \varphi, \quad x \in \Gamma_1. \end{aligned} \tag{5}$$

– Step 3: Compute the new approximation $\xi^{(k+1)}$ on Γ_2

$$\xi^{(k+1)} = v^{(k)}, \quad x \in \Gamma_2. \tag{6}$$

Remarks:

– The above iterative schemes has transformed the problem under consideration into two elliptic boundary value problem with weak mixed conditions (4)–(5)
– Convergence of the method depends on the convergence of iterative schemes (6)
– Scheme (6) is equivalent to scheme $\xi^{(k+1)} = (1-\tau)\,\xi^k + \tau v^k$, where τ is iterative parameter. This is two-layer iterative scheme [8].

$$\frac{\xi^{(k+1)} - \xi^{(k)}}{\tau} + B\xi^{(k)} = \Phi. \tag{7}$$

3.2 Approach to Determine $\dfrac{\partial U}{\partial n}|_{\Gamma_2}$

Let $\eta = \dfrac{\partial u}{\partial n}|_{\Gamma_2}$, we consider the following iterative schemes for finding η

Iterative Schemes QH2: With $\eta^{(0)} = 0, \forall k = 0, 1, 2, ...$ we solve the two problems respectively

– Step 1: Solve the problem with $u^{(k)}$

$$\begin{aligned} \Delta u^{(k)} - c^2 u^{(k)} &= f(x), \quad x \in \Omega, \\ \frac{\partial u^{(k)}}{\partial n} &= \eta^{(k)}, \quad x \in \Gamma_2, \\ u^{(k)} &= \varphi, \quad x \in \Gamma_1. \end{aligned} \tag{8}$$

– Step 2: Solve the problem with $v^{(k)}$

$$\begin{aligned} \Delta v^{(k)} - c^2 v^{(k)} &= f(x), \quad x \in \Omega, \\ \frac{\partial v^{(k)}}{\partial n} &= g, \quad x \in \Gamma_1, \\ v^{(k)} &= u^{(k)}, \quad x \in \Gamma_2. \end{aligned} \tag{9}$$

– Step 3: Compute the new approximation $\eta^{(k+1)}$

$$\eta^{(k+1)} = \frac{\partial v^{(k)}}{\partial n}, \quad x \in \Gamma_2. \tag{10}$$

Remarks:

– Similarly, problems (8)–(9) are elliptic boundary value problems with weak mixed boundary conditions.

– Iterative scheme (10) is equivalent to scheme $\eta^{(k+1)} = (1-\tau)\,\eta^{(k)} + \tau\dfrac{\partial v^{(k)}}{\partial n}$, where τ is the parameter. This is two-layer iterative scheme

$$\frac{\eta^{(k+1)} - \eta^{(k)}}{\tau} + B\eta^{(k)} = F. \tag{11}$$

It is difficult to prove that schemes converge in operator theory. However, we can test the convergence by experimentation.

3.3 Combined Iterative Scheme

In iterative schemes QH1 and QH2, in order to solve the problems (5), (6) and (8), (9), these problems have to be problems with weak mixed boundary conditions, i.e. Γ_2 is on a smooth boundary, which means the junction point between Γ_1 and Γ_2 is not a singularity. In case if the junction point between Γ_1 and Γ_2 is a singularity, we will have a problem with strong mixed boundary conditions. In this case, we need to combine the above iterative scheme with domain decomposition method. In the following part, we will consider the problem with strong mixed boundary conditions.

Consider the ill-posed problem when the junction point of Γ_1 and Γ_2 is singularity. Use domain decomposition method [1] to decompose $\Omega = \Omega_1 \cup \Omega_2$ by Γ.

We set $u_1 = u|_{\Omega_1}, u_2 = u|_{\Omega_2}$, $v_1 = v|_{\Omega_1}, v_2 = v|_{\Omega_2}$, $\lambda_u = \dfrac{\partial u_1}{\partial n}|_{\Gamma}, \lambda_v = \dfrac{\partial v_1}{\partial n}|_{\Gamma}$. Then, the ill-posed problem will be solved by one of the two following iterative scheme.

Iterative Scheme QH3: Determine the value of the function on Γ_2. Set $\xi = u_2|_{\Gamma_2}$, with $\xi^{(0)} = 0$

– Step 1: Solve the problem with $u_1^{(k)}$ and $u_2^{(k)}$. With $\lambda_u^{(0)} = 0$, perform domain composition algorithm
Step 1.1: Solve the problem with $u_1^{(k)}$ on Ω_1

$$\begin{aligned} \Delta u_1^{(k)} - c^2 u_1^{(k)} &= f, \quad x \in \Omega_1, \\ \frac{\partial u_1^{(k)}}{\partial n} &= g, \quad x \in \Gamma_1, \\ \frac{\partial u_1^{(k)}}{\partial n} &= \lambda_u^{(k)}, \quad x \in \Gamma. \end{aligned} \tag{12}$$

Step 1.2: Solve the problem with $u_2^{(k)}$ on Ω_2

$$\begin{aligned} \Delta u_2^{(k)} - c^2 u_2^{(k)} &= f, \quad x \in \Omega_2, \\ u_2^{(k)} &= \xi^{(k)}, \quad x \in \Gamma_2, \\ u_2^{(k)} &= u_1^{(k)}, \quad x \in \Gamma. \end{aligned} \tag{13}$$

– Step 2: Solve the problem with $v_1^{(k)}$ and $v_2^{(k)}$. With $\lambda_v^{(0)}$, perform domain composition algorithm
Step 2.1: Solve the problem with $v_1^{(k)}$ on Ω_1

$$\begin{aligned} \Delta v_1^{(k)} - c^2 v_1^k &= f, \quad x \in \Omega_1, \\ \frac{\partial v_1^{(k)}}{\partial n} &= \lambda_v^{(k)}, \quad x \in \Gamma, \\ v_1^{(k)} &= \varphi, \quad x \in \Gamma_1. \end{aligned} \tag{14}$$

Step 2.2: Solve the problem with $v_2^{(k)}$ on Ω_2

$$\begin{aligned} \Delta v_2^{(k)} - c^2 v_2^{(k)} &= f, \quad x \in \Omega_2, \\ \frac{\partial v_2^{(k)}}{\partial n} &= \frac{\partial u_2^{(k)}}{\partial n}, \quad x \in \Gamma_2, \\ v_2^{(k)} &= v_1^{(k)}, \quad x \in \Gamma. \end{aligned} \tag{15}$$

– Step 3: Compute the new approximation $\xi^{(k+1)}, \lambda_u^{(k+1)}, \lambda_v^{(k+1)}$

$$\begin{aligned} \xi^{(k+1)} &= v_2^{(k)}, \quad x \in \Gamma_2, \\ \lambda_u^{(k+1)} &= (1-\tau)\lambda_u^{(k)} - \tau \frac{\partial u_2^{(k)}}{\partial n}, \quad x \in \Gamma, \\ \lambda_v^{(k+1)} &= (1-\tau)\lambda_v^{(k)} - \tau \frac{\partial v_2^{(k)}}{\partial n}, \quad x \in \Gamma. \end{aligned} \tag{16}$$

Iterative Scheme QH4: Determine the value of the function on Γ_2, let $\eta = \frac{\partial u_2}{\partial n}|_{\Gamma_2}$. With $\eta^{(0)} = 0$, perform domain composition algorithm

– Step 1: Solve the problem with $u_1^{(k)}$ and $u_2^{(k)}$, with $\lambda_u^{(0)} = 0$
Step 1.1: Solve the problem with $u_1^{(k)}$ on Ω_1

$$\begin{aligned} \Delta u_1^{(k)} - c^2 u_1^{(k)} &= f, \quad x \in \Omega_1, \\ \frac{\partial u_1^{(k)}}{\partial n} &= \lambda_u^{(k)}, \quad x \in \Gamma, \\ u_1^{(k)} &= \varphi. \quad x \in \Gamma_1. \end{aligned} \tag{17}$$

Step 1.2: Solve the problem with $u_2^{(k)}$ on Ω_2

$$\begin{aligned} \Delta u_2^{(k)} - c^2 u_2^{(k)} &= f, \quad x \in \Omega_2, \\ \frac{\partial u_2^{(k)}}{\partial n} &= \eta^{(k)}, \quad x \in \Gamma_2, \\ u_2^{(k)} &= u_1^{(k)}, \quad x \in \Gamma. \end{aligned} \tag{18}$$

– Step 2: Solve the problem with $v_1^{(k)}$ and $v_2^{(k)}$.
With $\lambda_v^{(0)} = 0$, perform domain composition algorithm
Step 2.1: Solve the problem with $v_1^{(k)}$ on Ω_1

$$\begin{aligned} \Delta v_1^{(k)} - c^2 v_1^{(k)} &= f, \quad x \in \Omega_1, \\ \frac{\partial v_1^{(k)}}{\partial n} &= \lambda_v^{(k)}, \quad x \in \Gamma, \\ v_1^{(k)} &= \varphi, \quad x \in \Gamma_1. \end{aligned} \tag{19}$$

Step 2.2: Solve the problem with $v_2^{(k)}$ on Ω_2

$$\begin{aligned} \Delta v_2^{(k)} - c^2 v_2^{(k)} &= f, \quad x \in \Omega_2, \\ \frac{\partial v_2^{(k)}}{\partial n} &= \frac{\partial u_2^{(k)}}{\partial n}, \quad x \in \Gamma_2, \\ v_2^{(k)} &= v_1^{(k)}, \quad x \in \Gamma. \end{aligned} \tag{20}$$

– Step 3: Compute the new approximation $\eta^{(k+1)}, \lambda_u^{(k+1)}, \lambda_v^{(k+1)}$

$$\begin{aligned} \eta^{(k+1)} &= \frac{\partial v_2^{(k)}}{\partial n}, \quad x \in \Gamma_2, \\ \lambda_u^{(k+1)} &= (1-\tau)\lambda_u^{(k)} - \tau \frac{\partial u_2^{(k)}}{\partial n}, \quad x \in \Gamma, \\ \lambda_v^{(k+1)} &= (1-\tau)\lambda_v^{(k)} - \tau \frac{\partial v_2^{(k)}}{\partial n}, \quad x \in \Gamma. \end{aligned} \tag{21}$$

The convergence of the method depends on the convergence of iterative schemes (16), (21), which is the iterative scheme in domain decomposition method with strong mixed boundary value problems. The convergence is proposed in [1].

4 Numerical Examples

To test the convergence as well as convergence rate of the proposed iterative schemes, we use difference schemes and formulas of second order of accuracy to transform differentiation problems with weak mixed boundary condition in a rectangular domain into a vector three-point equation, then we apply the method of complete reduction [7] to solve systems of equations to find approximate solutions of sub-problems, from which the solution of the original problem is found. Functions of numerical solutions were developed in the library RC 2009 [9]. In the experimental part, domain Ω is the rectangular domain,

$\Omega = \{0 \leqslant x \leqslant a, 0 \leqslant y \leqslant b\}$ is covered by an uniform grid $M \times N = 64 \times 64$ nodes. Computational error $\varepsilon = \max\limits_{(i,j)} \left\{ \left\| u^{(k+1)} - u^{(k)} \right\| \right\}$, if exact solutions are known, errors after each iteration is $\varepsilon = \max\limits_{(i,j)} \left\{ \left\| u^{(*)} - u^{(k)} \right\| \right\}$ where $u^{(*)}$ denotes the correct solution of the problem.

4.1 Test Results QH1 and QH2

Consider the problem

$$\Delta u - c^2 u = f, \quad (x, y) \in [0, a] \times [0, b],$$
$$\frac{\partial u}{\partial n} = g, \quad u = \varphi \quad x \in \Gamma_1,$$

where $\Gamma_1 = \{0 \leqslant x \leqslant a; y = 0, b\} \cup \{x = 0, a; 0 \leqslant y \leqslant b\}$; $\Gamma_2 = \{0 \leqslant x \leqslant a; y = b\}$. We will find solutions of the problem by iterative schemes QH1 and QH2 described in (4)–(6) and (8)–(10).

Case 1: Given exact solution $u^*(x, y) = \sin x \sin y$. Based on the exact solution, we will determine the right-hand side function $f(x)$, boundary condition functions $\varphi(x), g(x)$ so that the error between correct solutions and approximate solutions is determined according to the algorithm. The results of computation comparison of exact solutions and approximate solutions are given in Table 1.

Table 1. Comparison of exact solutions and approximate solutions

Number of iteration	Error QH1	Error QH2
50	0.0193	0.0073
100	0.0114	0.0042
200	0.0060	0.0022
500	0.0023	8.10^{-4}

Case 2. Unknown exact solution, $f = x^2 - y + \sin x \sin y$; $\varphi = \cos x \sin y$; $g = x \cos y$, so that error between two successive iterations is determined according to the algorithm. The results of computation comparison of two successive iterations are given in Table 2 and approximate solution graph are given in Fig. 1.

Table 2. Comparison of two successive iterations

Number of iteration	Error QH1	Error QH2
50	0.0017	0.0013
100	0.0010	9.10^{-4}
200	6.10^{-4}	4.10^{-4}
500	3.10^{-4}	1.10^{-4}

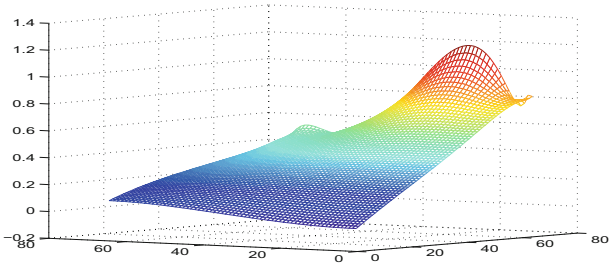


Fig. 1. Approximate solution graph

4.2 Test Results of QH3 and QH4

Consider the problem

$$\Delta u - c^2 u = f, \quad (x, y) \in [0, 2a] \times [0, b],$$
$$\frac{\partial u}{\partial n} = g, \quad u = \varphi, \quad x \in \Gamma_1,$$

where $\Gamma_1 = \{0 \leqslant x \leqslant 2a, y = 0\} \cup \{0 \leqslant x \leqslant a; y = b\} \cup \{x = 0, 2a; 0 \leqslant y \leqslant b\}$, $\Gamma_2 = \{a \leqslant x \leqslant 2a; y = b\}$. We find solutions of the problem by iterative schemes QH3 and QH4 described in (12)–(16) and (17)–(21).

Case 1. Given exact solution $u^*(x, y) = \sin x \sin y$. The results of computation comparison of exact solutions and approximate solutions are given in Table 3.

Table 3. Comparison of exact solutions and approximate solutions

Number of iterations	Error QH3	Error QH4
50	0.0423	0.0165
100	0.0361	0.0105
200	0.0141	0.0054
500	0.0035	0.0013

Case 2. Unknown exact solutions, $f = x^2 - y + \sin x \sin y; \varphi = \cos x \sin y; g = x \cos y$. The results of computation comparison of two successive iterations are given in Table 4 and approximate solution graph are given in Fig. 2.

Table 4. Comparison of two successive iterations

Number of iterations	Error QH3	Error QH4
50	0.0493	0.0173
100	0.0214	0.0092
200	0.0090	0.0052
500	0.0042	0.0015

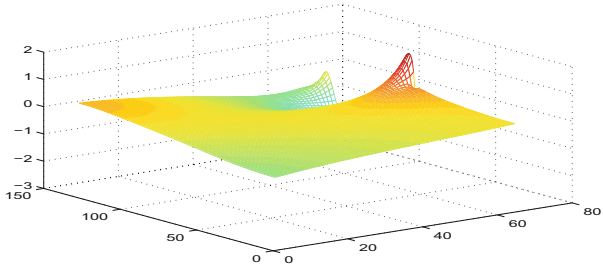


Fig. 2. Approximate solution graph

5 Concluding Remarks

From the numerical results presented in this study, it can be seen that the proposed iterative scheme converges. To increase the convergence rate we select ideal iterative parameters.

In the two methods, updates derivative has better convergence rate than updates function method.

The previous iterative schemes can be expanded for problems with more complex boundary conditions.

Acknowledgment. This work is supported by Thai Nguyen University under the project DH2016-TN07-02.

References

1. Quang, A.D., Quang, V.V.: Domain decomposition method for solving strongly mixed boundary conditions. J. Comput. Sci. Cybern. **22**(4), 307–318 (2006)
2. Kozlov, V.A., Mazya, V.G., Fomin, A.V.: An iterative method for solving the Cauchy problem for elliptic equations. Comput. Math. Math. Phys. **31**, 45–52 (1991)
3. Kraus, A.D., Aziz, A., Welty, J.: Extended Surface Heat Transfer. Wiley, New York (2001)
4. Marin, L.: An alternating iterative MFS algorithm for the Cauchy problem for the modified Helmholtz equation. Comput. Mech. **45**, 665–677 (2010)
5. Marin, L.: Boundary element-minimal error method for the Cauchy problem associated with Helmholtz-type equations. Comput. Mech. **44**, 205–219 (2009a)

6. Marin, L.: An iterative MFS algorithm for the Cauchy problem associated with the Laplace equation. Comput. Model Eng. Sci. **48**, 121–153 (2009b)
7. Samarskij, A., Nikolaev, E.: Numerical Methods for Grid Equations, vol. 2. Birkhauser, Basel (1989)
8. Samarskij, A.: The Theory of Difference Schemes. Marcel Dekker, New York (2001)
9. Hai, T.H., Quang, V.V.: Results of the RC2009 program on solving some elliptic boundary problems with constant coefficients. J. Sci. Technol. Thai Nguyen Univ. **69**(07), 56–63 (2010)

BKCA, an E-Consultancy System for Studying

Thi Thu Trang Nguyen(✉), Viet Thang Dinh, and Tuan Dung Cao

School of Information and Communication Technology, Hanoi University of Science and Technology, 1 Dai Co Viet, Hai Ba Trung, Hanoi, Vietnam
{trangntt,dungct}@soict.hust.edu.vn

Abstract. In education organizations, a number of quotidian and duplicated consulting questions for training programs, university regulations... make difficulties, inconvenient and time-consuming for consultants. This work proposes a method to suggest answers of similar questions, which may help both students and consultants in orientation. This method includes two main tasks: (i) query generation which extracts key phrases from input questions, (ii) searching which find out most similar questions with their answers and relevant scores using extracted key phrases. For the first task, a check strategy was used to build key phrase candidates and the Naïve Bayes Classifier was used to select key phrases. The two different approaches were experimented for the second one: (i) Similarity Comparison Searching, and (ii) Sorl Search Engine. The precision of the key phrase extraction is about 69 %. BKCA, an e-consultancy system for SoICT-HUST, was built and experimented with about 46 % relevant question-answer pairs for the Similarity Comparison searching approach, and about 39 % for the Sorl Search Engine one.

Keywords: Keyphrase extraction · Recommendation system · Naïve Bayes · E-consultance

1 Introduction

Giving advices for students about studying, regulation... via emails is an important task in universities. This is a convenient and time-saving way for students who need consultancies. With a number of students and hence questions, consultants could not reply all students' questions promptly due to time limitation. They also need to refer to university regulations, old questions (via email) or other information for the answers. Different consultants may receive the same or similar questions at different times. This raises a need of a system, which not only stores and processes question-answer pairs but also could suggest a list of similar questions (and their answers) to the input question from previous question-answer pairs. BKCA, Bach-Khoa Consultancy Application, is such a system, which supports not only for consultants but also for students.

For the problem of giving a list of similar questions for an input question, there are two main tasks: (i) *Query Generation* that extracts key phrases from the input question and (ii) *Searching* that use generated key phrases as a query

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_20

to find out relevant questions from historical data (i.e. questions-answer pairs). Key phrases are informative ones that represent the main content of a document.

For the first task, there are a number of researches on key phrase extraction in English either using supervised or unsupervised approach. For example, *Feifan Liu* and his colleagues [4] used unsupervised way to make a query, but his work focused on how to find key words, but not the meaning of them. By using a supervised system, Witten and his colleagues [5] proposed a supervised learning machine for extracting key words with three features of frequency, inverse frequency and distance where key word appears. However, this work did not show the combination of key words, i.e. key phrases. Lee and Croft [7] came up with the idea of using two ways together (e.g. semi-supervised): (i) Find key words based on documents which are chosen by users, and (ii) Choosing candidates by dividing documents into chunks then use a statistical approach and machine learning to define key words. All these works were devoted for English, but not for Vietnamese.

An effort for Vietnamese proposed a method to generate key words by sending terms of a specific-domain ontology to the Google search engine [8]. However, to build ontology is a complex task with much time and expert knowledge. Vietnamese has a complex grammar structure and multi-meaning words. Therefore, it is a big challenge to make key phrase extraction and combine them. A research team of *Phạm Thị Thu Uyên*, which is used in automatic inquiry system, combined snowball method of Agichtein and Gravano [3] and searching machine method of Ravichandran Hovy [6] to extract model of meaning relation in Vietnamese documents. However, finding key words is not enough, sometimes the combination of key words make it change its own meaning. Therefore, the combination of extracted key words plays an important role and improve the information query.

For the second task, there are a number of researches and tools for the query tasks such as indexing, processing, scoring and ranking the result. This problem can be done by using a search engine such as Google Engine, Solr search, Elastic search... However, these systems do not have any support for the Vietnamese language.

In this paper, we propose an unsupervised model combined with machine learning and statistics for the key phrase extraction. There are three tasks in preprocessing Vietnamese documents: (i) Pre-process of the data because the data is raw data, (ii) Key phrase extraction: Candidate phrases are chosen by observation and experiment. (iii) Searching and Ranking to find the question-answer pairs, which are related to the input question. Two searching approaches were experimented: (i) Search Engine Solr, and (ii) Similarity Comparison.

The rest of this paper is organized as follows. Section 2 presents the proposed model for an e-consultancy system for studying – BKCA. In Sects. 3 and 4, the solution for the two main tasks, i.e. Key phrase extraction and Searching, is described. Section 5 is the experiment for the BKCA system. The final section gives conclusions and presents future works.

2 Proposed Model

Figure 1 illustrates our proposed model for suggesting similar questions to the input question. The input question is firstly pre-processed to clean data, normalize acronyms, phrase extraction, word segmentation and POS tagging. Key phrases are then extracted to build query for the next step. List of similar questions with their answers are finally found using the generated query. These three tasks are presented in below subsections.

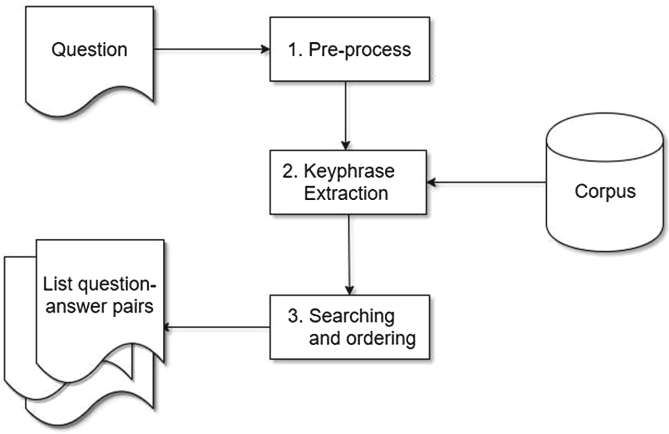


Fig. 1. Proposed model for suggesting similar questions.

2.1 Data Pre-processing

In this task, the raw data from the consultancy emails are collected and pre-processed. This includes the following sub-tasks:

1. *Question-answer extraction*: Dividing the email into question-answer pairs since it might have many question in an email, if in the training phase;
2. *Normalization*: Normalizing the data using an Acronym Dictionary, such as: "chương trình đào tạo" *(curriculum)* for "ctdt" or "CTĐT". Beside an acronym dictionary, we also used some rules for recognizing abbreviations such as all capitalization, all consonants (may be in lowercase), both characters and number...;
3. *Phrase extraction*: Dividing the input question into phases using punctuations;
4. *Word segmentation*: Segmenting words for each phrases;
5. *POS tagging*: Give POS (Part-of-Speech) for all word in phrases

Table 1 gives an example for all pre-processing steps for a question. We adopted the JVnTextPro [2] for the two last steps.

Table 1. An example of pre-processing task

Question		Xin cho em hỏi, theo ctđt, lớp có số lượng đăng ký bao nhiêu thì bị hủy lớp?? *(Could you please let me know, based on the syllabus, a class with how many registrations will be cancelled?)*
Sub-task		**Result**
(i)	Question-answer extraction	(In training phase)
(ii)	Normalization	ctđt: chương trình đào tạo
(iii)	Phrase extraction	Xin cho em hỏi theo chuong trình đào tạo lớp có số lượng đăng ký bao nhiêu thì bị hủy lớp?
(iv)	Word segmentation	Xin cho em hỏi theo chương_trình đào_tạo lớp có số_lượng đăng_ký bao_nhiêu thì bị hủy lớp?
(v)	POS tagging	Xin/V cho/V em/P hỏi/V Theo/V chương_trình/N đào_tạo/N lớp/N có/V số_lượng/N đăng_ký/N bao_nhiêu/X thì/V bị/V hủy/V lớp/N ạ/X?

2.2 Key Phrase Extraction

To automatically predict key phrases from the input question, both supervised and unsupervised approaches were used. Key words were firstly automatically extracted using Naïve Bayes classification and then an unsupervised approach for combining key words into key phrases. In the training corpus for the Naïve Bayes classification model, key phrases were semi-automatically selected by expert opinions. For example, three key phrases were found: "chương trình đào tạo" *(curriculum)*, "số lượng đăng ký" *(number of registration)*, "hủy lớp" *(cancel class)* for the question in Table 1.

2.3 Searching and Ranking

The purpose of this task is to find out the most similar questions to the input question from the historical question-answer pairs. As presented, extracted key phrases are used to search and rank the result. In this step, we experimented two approaches: (i) using Solr Search Engine with a plug-in for Vietnamese language, (ii) using the Similarity Comparison method.

3 Key Phrase Extraction

The model of key phrase extraction includes two stages: Training stage and Extraction stage, which describes in Fig. 2. Key words are firstly identified using POS information. These words are then combined to build key phrase candidates by some unsupervised rules. Key phrases may be a key word or a combination

of key words. In the training stage, following the above process, key phrases are chosen for all questions in the training corpus by experts.

Candidate identification and combination. Based on observations on expert opinions, verbs, nouns, adjectives may have important meaning in sentences and hence are considered as candidate key words. Due to their ambiguity, unknown words are also listed as candidate key words. Since Vietnamese has a complex grammar structure and multi-meaning words, key words are not enough. Therefore, adjacent key words are combined to build candidate key phrases, which may increase the accuracy and reliability for Vietnamese. In this case, the two combining rules are used: (i) if there is no adjacent key words to the observing one, this key word is chosen for the next step, and (ii) if the key word c_1 stands next to the key word c_2, a check function is then calculated for their combination c_1c_2 in the corpus, as illustrated in Eq. 1.

$$\text{check} = \frac{\text{count}(c_1c_2)}{\text{count}(\text{Nonc}_1c_2) + \text{count}(c_1c_2)} \tag{1}$$

where

- $\text{count}(c_1c_2)$: number of questions which has the combination c_1c_2 (i.e. c_2 stands next to c_1)
- $\text{count}(\text{Nonc}_1c_2)$: number of questions which have the two key words c_1 and c_2 but not the combination c_1c_2.

If the value of the check function is equal or greater than 50 % then c_1c_2 is considered as a candidate key phrase. If the value of check is less than 50 % then c_1 and c_2 are added to the key phrase list.

Key phrase selection. To predict and reduce unnecessary key phrases, a supervised method is used with a Naïve Bayes classification. After having some observations and experiments, we finally found five optimal features that can be used to decide the key phrases that can represent a question.

Table 2 shows these five optimal features. The first feature (F_1) is the frequency level of the candidate that appears in the input question. The second feature (F_2) is the percentage level of the question that contains the candidate. The third and the fourth one (F_3, F_4) let us know if the candidate appears in the subject or the in answer or not. The last feature (F_5) represents the distance level of the current candidate to the beginning of the questions, i.e. number of key phrases standing before the current candidate. After observing data, the level of F_1, F_2 and F_5 are proposed in the Fig. 2. They are relative level, which based on their minimum, average and maximum values.

The above five features are used to trained with the Naïve Bayes model due to its simplicity and efficiency. This model is then used to classify if the candidate is a key phrase or not for the input question. The query is finally built based on the list of key phrases.

Table 2. Optimal features for key phrase prediction

Feature	Description	Possible values
$\mathbf{F}_1$	The frequency of the candidate which appears in the input question	Very high, High, Medium, Low, Very low
$\mathbf{F}_2$	The percentage of the question that contains the candidate	Very high, High, Medium, Low, Very low
$\mathbf{F}_3$	Does candidate from the question appear in the subject?	Yes, No
$\mathbf{F}_4$	Does candidate appear in the answer?	Yes, No
$\mathbf{F}_5$	Distance: number of key phrases standing before current candidate	Very high, High, Medium, Low, Very low

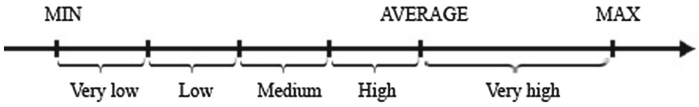


Fig. 2. Levels of feature values.

4 Searching and Ranking

To suggest a list of similar questions and their answers, we experimented 2 approaches: (i) Similarity comparison, and (ii) Sorl Search engine with a Vietnamese plug-in.

4.1 Similarity Comparison

Figure 3 describes the Searching method based on Similarity Comparison. The input is a query, which is a list of key phrases {kp_1, kp_2, ... kp_n} (n<10). The output is top 10 question-answer pairs, which are relevant to the input question. There are three steps in this approach: (i) key phrase comparison, (ii) similarity calculation, and (iii) ranking.

Firstly, for each predicted key phrase in the input question, identifications (ID) of questions having that key phrase are found. The number of times that the question has a key phrase is also recorded. For example, for a query having {đăng_ký, học_tập, cử_nhân_công_nghệ, đồ_án} *(register, study, bachelor of IT, thesis)*, question ids are listed as "1, 4, 5, 6, 8, 9, 10, 17, 39", in detail:

– "đăng_ký" *(register)* appears in question ID = 1, 17, 4, 5, 8
– "học_tập" *(study)* appears in question ID = 1, 4, 5, 9
– "cử_nhân_công_nghệ" *(bachelor of IT)* appears in question ID = 1, 17, 39, 5, 10
– "đồ_án" appears *(thesis)* in question ID = 39, 4, 5, 6.

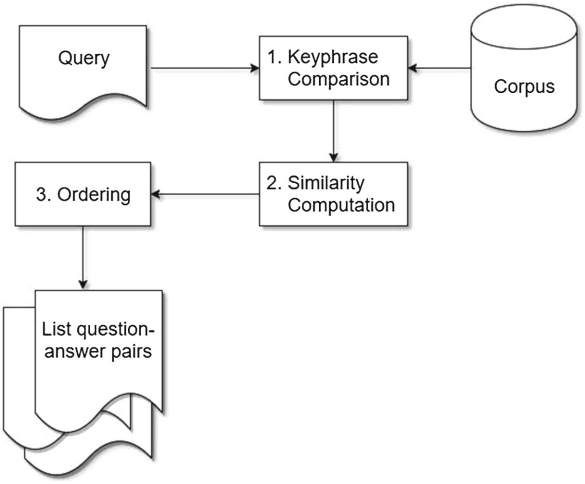


Fig. 3. Similarity comparison model.

The question list is ordered based on the number of times that the question has a key phrase, e.g. "ID = 5" appears 4 times; "ID = 1" appears 3 times; "ID = 4" appears 3 times; "ID = 39" appears 2 times; "ID = 17" appears 2 times. The question having the highest times is considered as the most similar to the input question. If there is more than one question having the same time, the weights of these questions are considered. The weight of a question is decided by the number of true key phrases it has, and then if it is a newer question. For example, the questions "1" and "4" both have 3 times, but the question "1" has more priority since it has a true key phrase "cử_nhân_công_nghệ" *(bachelor of IT)* while the question "4" not. The question "39" is ranked before the question "17" since it is newer, then has more updated information.

4.2 Solr Search Engine with a Vietnamese Plug-in

We experimented another approach to search similar questions based on a query, i.e. using a search engine – Apache Solr search. Apache Solr [1] is an open source search engine that allows users customized and adjusted according to the user's goals. Solr is built Apache Lucene based library support full-text search performance advanced. Currently, Solr search has supported for multiple languages such as English, French, and Japanese... without supporting for the Vietnamese language.

Therefore, to keep the meaning of Vietnamese words, we built a plugin that tokenizes Vietnamese words to the Apache Solr. Some other main tasks to search similar questions based on a query are (i) Removal of Vietnamese stop words, (ii) Synonym Filtering, (iii) Removal of word delimiters, and (iv) Lowercase processing. Word delimiters can be punctuations, or some special rules for tokenizing some special words, such as "BachKhoa" was token to "Bach" "Khoa".

All these words were then all converted to lowercase and finally indexed for searching using the Apache Solr search engine.

All parts of question-answer emails (subject, question, answer) were indexed with different weights, which were chosen based on the observation and experiment (Table 3). The Subject part has the highest priority, while Answer one has the lowest weight.

Table 3. Optimal weight of email parts

Email part	Subject	Question	Answer
Weight	20	10	5

5 Experiment

This section describes how we experimented our proposal for the e-consultancy system, BKCA. This system is being deployed for the School of Information and Communication Technology (SoICT), Hanoi University of Science and Technology (HUST).

5.1 Preparation of Corpus

We investigated on the raw data including more than 1,000 previous emails between students and consultants at SoICT-HUST from April 2015 to April 2016 (about 1 year). We found a number of duplicated questions or answers with out-of-date regulations in the raw data. About one third of emails had threaded conversations between students and consultants, which might give several question-answer pairs per email. Finally, about 620 question-answer pairs were semi-automatically. We divided in to 3 parts: 70 % for training, 20 % for optimization, and 10 % for test.

5.2 Evaluation of Key Phrase Extraction

To evaluate the key phrase extraction step, Precision is calculated as the rate of key phrases that are classified correctly while Recall is the rate of correct predicted key phrases over the actual key phrases. Table 4 shows the result of the key phrase extraction step using Naïve Bayes with five optimal features. The Precision, Recall and F-score are nearly 69 %, which is good for this problem, compared to other languages.

5.3 Evaluation of Searching and Ranking

For each question in 10 % questions in the test corpus, the BKCA system suggested a list of similar questions to the input question. The expert then checked

Table 4. Evaluation of key phrase extraction using Naïve Bayes

Measure	Precision	Recall	F-score
Rate (%)	68.8 %	68.9 %	68.8 %

out all truly relevant questions in that suggesting list. Table 5 shows the experiment results of the two searching approaches. The Similarity Comparison approach gives a better result (about 6 %) and less time-consuming than the Solr search engine one. The Solr search engine also depends on the quality and performance of the Vietnamese tokenizer, while the Similarity Comparison depends on only the extracted key phrases.

Table 5. Comparison between the two searching approaches

Searching approach	Relevant rate
Solr Search engine	38.4 %
Similarity Comparison	46.0 %

6 Conclusion

This study puts an initial step for the problem of suggesting similar questions in Vietnamese with two main tasks: (i) Key phrase extraction to build a query for each question, and (ii) Searching and ranking similar questions based on the generated query. In key phrase extraction, a semi-supervised approach was used: (i) POS information and combination rules were used to build candidate key phrases from key words, and (ii) Naïve Bayes classification with five optimal features was used to choose key phrases from candidates. The five features are: (i) F_1: The frequency level of the candidate in the question (in the corpus) (ii) F_2: The percentage level of the question (in the corpus) having the candidate, (iii) F_3: If the candidate appears in the subject? (iv) F_4: If the candidate appears in the answer, and (v) F_5: the distance level from the candidate to the beginning of the question. Using classified method, Naïve Bayes classification with database of 620 question-answer pairs brought a result with 68.8 % corrected key word extraction. We experimented two approaches to search and rank similar questions: (i) Similarity Comparison, and (ii) Sorl search engine with a Vietnamese plug-in. The Similarity Comparison approach gives a better result (about 6 %) and less time-consuming than the Solr search engine one. About 46 % questions were marked as relevant questions-answers to the input question. In the future, we will study more on how to analyze the meaning of key words and the question/answer. The system is being deployed in the way to get feed back from users to improve data and model.

Acknowledgments. This research is funded by Hanoi University of Science and Technology through the "E-consultancy system for studying in SoICT-HUST" project – grant number T2016-PC-049.

References

1. Apache Solr Search. http://lucene.apache.org/solr/
2. Nguyen, C.T., Phan, X.H., Nguyen, T.-T.: JVnTextPro: a Java-based Vietnamese Text Processing Tool (2010). http://jvntextpro.sourceforge.net/
3. Agichtein, E., Gravano, L.: Snowball: extracting relations from large plain-text collections. In: Proceedings of the Fifth ACM Conference on Digital Libraries, DL 2000, Texas, USA (2000)
4. Liu, F., Pennell, D., Liu, F., Liu, Y.: Unsupervised approaches for automatic keyword extraction using meeting transcripts. In: Proceeding of the Conference of the North American Chapter of the ACL, Boulder, Colorado (2009)
5. Witten, I.H., Paynter, G.W., Frank, E., Gutwin, C., Nevill-Manning, C.G.: KEA: practical automatic key phrase extraction. In: Proceedings of the Fourth ACM Conference on Digital Libraries, California, USA (1999)
6. Ravichandran, D., Hovy, E.: Learning surface text patterns for a question answering system. In: Proceedings of the 40th Annual Meeting on Association for Computational Linguistics (ACL 2002), Philadelphia, Pennsylvania (2002)
7. Lee, C.-J., Bruce Croft, W.: Generating queries from user-selected text. In: Proceedings of 4th Information Interaction in Context Symposium (IIIX 2012), Nijmegen, The Netherlands (2012)
8. Cao, T.-D., Dieng-Kuntz, R., Fiès, B.: An ontology-guided annotation system for technology monitoring. In: IADIS International WWW/Internet 2004 Conference, Madrid, Spain, 6–9 October 2004

Classifying Human Body Postures by a Support Vector Machine with Two Simple Features

Nguyen Van Tao, Nong Thi Hoa(✉), and Quach Xuan Truong

Thai Nguyen University of Information and Communication Technology,
Quyet Thang, Thai Nguyen, Vietnam
{nvtao,nthoa,qxtruong}@ictu.edu.vn

Abstract. Human behaviour analysis helps to monitor a person's daily activities and detect home care emergencies. Classifying posture is an important step of human behaviour analysis. Many studies improve the accuracy of classifying. However, the number of features is big or extracting these features uses complicated formulas. Therefore, we proposed two features with simple computing. Two new features are formulated from the height and the square showing the human body's silhouettes. Then, we choose a non-linear Support Vector Machine to classify postures based on proposed features. Experiments show Support Vector Machine classify effectively and better than other methods.

Keywords: Posture classification · Human behaviour analysis · Support Vector Machine

1 Introduction

A visual surveillance system attempts to detect, recognize, and track some objects from camera. Intelligent visual surveillance is developed to increase understanding and description of human behaviours. Indoor behaviour surveillance can automatically monitor a person's daily activities and detect home care emergencies. A surveillance system usually includes fours steps: environment modelling, person detection, feature extraction, behaviour prediction. Classifying postures is an important stage of the behaviour prediction to make a warning alert or nothing.

Many studies on proposing effective features for available classifiers have been developed for human posture classification. Chia-Feng Juang et al. [1] presented methods locating the head, the center of the body, the tips of the feet, the tips of the hands, the elbows, and the knees of the human body. Then, using matching methods based on Kalman filtering algorithm to classify postures. Moreover, the 3D human skeleton was reconstructed from set of significant points of 2D images. Nacira Zerrouki and Amrane Houacine [11] proposed a Multilayer Perceptron network with back propagation learning algorithm classifying based on Singular Value Decomposition coefficients and the height-width ratio of human body's contour. Jamie Shotton et al. [8] predicted 3D positions of body joints to

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_21

design an intermediate body parts representation. Then, Randomized decision forests were implemented to classify. Chia-Feng Juang et al. [2] used a Gaussian-kernel-based Support Vector Machine based on projections and length-width ratio of the human body's silhouette. Vincent Girondel et al. [14] extracted two features from face and hands localization for belief theory-based classifiers. Chen Wu and Hamid Aghajan [3] used color distribution to find fitted ellipses corresponding to the human body. Then, a Gaussian Mixture Model was applied for classifying. Additionally, a 3D skeleton was reconstructed from fitted ellipses. Rita Cucchiara et al. [12] used features based on the blob's silhouette to create probabilistic projection maps for a Bayesian classifier. Lutz Goldmann et al. [9] used the contour-based shape descriptor (MPEG-7) and the projection histogram as features. Classifying was carried out by a Gaussian mixture model and the k-nearest neighbour method. Isaac Cohen and Hongxia Li [7] introduce a 3D human shape description from human body's silhouettes. Then, using a support vector machine for inferring postures.

Generally, formulating features are complicated because of complex equations [2,11]. As a result, the computational complexity is large and computing time increases. In this paper, we propose two simple features based on counting pixels of the human body's silhouette in only one time. Support Vector Machine is an effective method for classifying. Therefore, we choose a non-linear Support Vector Machine to classify postures in order to improve the ability of classifying. We hope to decrease time of computing and still classify effectively. Experiments are conducted to prove the effectiveness of new features and choosing a Support Vector Machine classifier.

The paper is structured as follows. Support Vector Machine is described in Sect. 2. Section 3 introduces feature extraction from the human body's silhouette. In Sect. 4, we present experimental results.

2 Support Vector Machine

2.1 Definition

A Support Vector Machine (SVM) is a supervised learning technique of machine learning to apply for both classification and regression.

Given a set of training examples, each example belongs to one of two categories. A SVM training algorithm builds a model that assigns new examples into one category [4]. A SVM model is a representation of the examples as points in space and maps the examples into categories by a clear gap.

New examples are predicted to belong to a category based on which side of the gap they fall on. Moreover, SVM can efficiently perform a non-linear classification by using the kernel trick.

2.2 Linear SVM

Given a training dataset of n points of the form $(\overrightarrow{x_1}, y_1), ..., (\overrightarrow{x_n}, y_n)$ where the y_i are either 1 or -1, each indicating the class to which the point $\overrightarrow{x_i}$ belongs to. Each $\overrightarrow{x_i}$ is a p-dimensional real vector.

Find the "maximum-margin hyper-plane" that distance between the hyper-plane and the nearest point $\overrightarrow{x_i}$ from either group is maximized.

Any hyper-plane can be written as the set of points $\overrightarrow{x}$ satisfying

$$\overrightarrow{w}.\overrightarrow{x} - b = 0 \tag{1}$$

where $\overrightarrow{w}$ is the normal vector to the hyper-plane. The parameter $\frac{b}{\|\overrightarrow{w}\|}$determines the offset of the hyper-plane from the origin along the normal vector $\overrightarrow{w}$.

2.3 Non-linear Classification

The kernel trick is applied to create non-linear classifiers. The resulting algorithm is formally similar, except that every dot product is replaced by a non-linear kernel function. Some common kernels include:

Polynomial (homogeneous)

$$k(\overrightarrow{x_i}, \overrightarrow{x_j}) = (\overrightarrow{x_i}.\overrightarrow{x_j})^d \tag{2}$$

Polynomial (inhomogeneous)

$$k(\overrightarrow{x_i}, \overrightarrow{x_j}) = (\overrightarrow{x_i}.\overrightarrow{x_j} + 1)^d \tag{3}$$

Gaussian radial basis function

$$k(\overrightarrow{x_i}, \overrightarrow{x_j}) = exp(-\gamma\|\overrightarrow{x_i} - \overrightarrow{x_j}\|)^2) \tag{4}$$

for $\gamma \geq 0$.

Hyperbolic tangent:

$$k(\overrightarrow{x_i}, \overrightarrow{x_j}) = tanh(h\overrightarrow{x_i}.\overrightarrow{x_j} + c), \tag{5}$$

for $h > 0$ and $c < 0$.

2.4 Computing the SVM Classifier

A SVM learns a non-linear classification rule which corresponds to a linear classification rule for the transformed data points $\varphi(\overrightarrow{x_i})$. Moreover, using a kernel function k need satisfies

$$k(\overrightarrow{x_i}, \overrightarrow{x_j}) = \varphi(\overrightarrow{x_i}).\varphi(\overrightarrow{x_j}) \tag{6}$$

The classification vector $\overrightarrow{w}$ is computed by

$$\overrightarrow{w} = \sum_{i=1}^{n} c_i y_i \varphi(\overrightarrow{x_i}) \tag{7}$$

where the c_i are obtained by solving the optimization problem

$$\max f(c_1, ..., c_n) = \sum_{i=1}^{n} c_i - \sum_{i=1}^{n} \sum_{j=1}^{n} y_i c_i k(\overrightarrow{x_i}, \overrightarrow{x_j}) c_j y_j \tag{8}$$

subject to

$$\sum_{i=1}^{n} c_i y_i = 0 \tag{9}$$

and

$$0 \leq c_i \leq \frac{1}{2n\lambda}, \forall i \tag{10}$$

Finally, new points are classified by

$$\overrightarrow{z} \rightarrow sgn(\sum_{i=1}^{n} c_i y_i k(\overrightarrow{x_i}, \overrightarrow{z}) + b) \tag{11}$$

3 Human Body Feature Extractions

3.1 Two New Features

The human body is segmented by a background subtraction method. These images are taken indoors and from a single fixed camera. Two simple features are extracted from the human body's silhouette including Tc and Tv. Tc is the height ratio of person and image. Tv is the pixel ratio of a region showing the detail part of the human body and the complete human body. Using the ratio of values avoids the body size's difference when images are taken from various distances.

For counting Tc, locate the top of the body TB and the bottom of the body BB. Tc is formulated by

$$Tc = \frac{BB - TB}{N} \tag{12}$$

where N is the height of the current image.

Figure 1 shows detail values of Tc.

Tv is computed based on the square showing human body. First, formulate the square of the top part of body (PB). Then, the square of full body (FB) is summarized. The square is equal to the number of pixels fulfilling a region of human body. Variable h presents the height of the body's top part. PB is the number of pixels in the region from the top to the h^{th} row. The equation of Tv is shown

$$Tv = \frac{PB}{FB} \tag{13}$$

Figure 2 presents how to formulate values of Tv. White lines show regions using for Eq. (13).

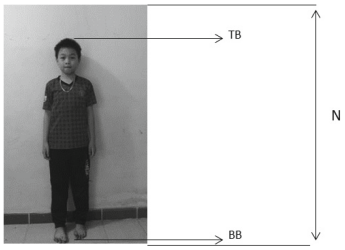


Fig. 1. Calculation of Tc

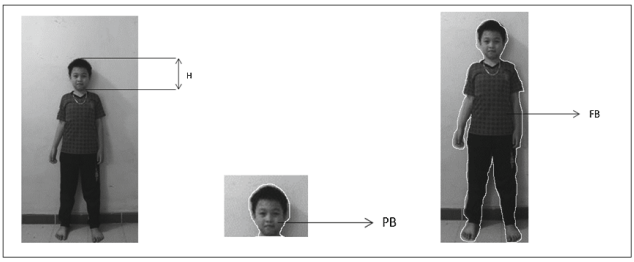


Fig. 2. Calculation of Tv

3.2 Comparing to Other Studies

We summarize the number of features and the description of these features. Results are presented in Table 1. Feature Description in Table 1 show our approach is more simple than other studies.

Table 1. The number of features and feature description

Methods	# Feature	Feature description
N. Zerrouki [11]	42	Singular Value Decomposition coefficients
L. Goldmann [9]	maximum 64	The contour-based shape descriptor (MPEG-7)
C. Juang [2]	84	Fourier transform coefficients
V. Girondel [14]	4	Face and hands Location
Our approach	2	Number of pixels on human body's silhouette

4 Experiment

4.1 Dataset

Our data consist of 740 images including 60 images of the training set and 680 images of the testing set. The size of images is 368×490. Figure 3 shows typical training images of four postures including sitting, lying, bending, and standing.



Fig. 3. Examples of the training set

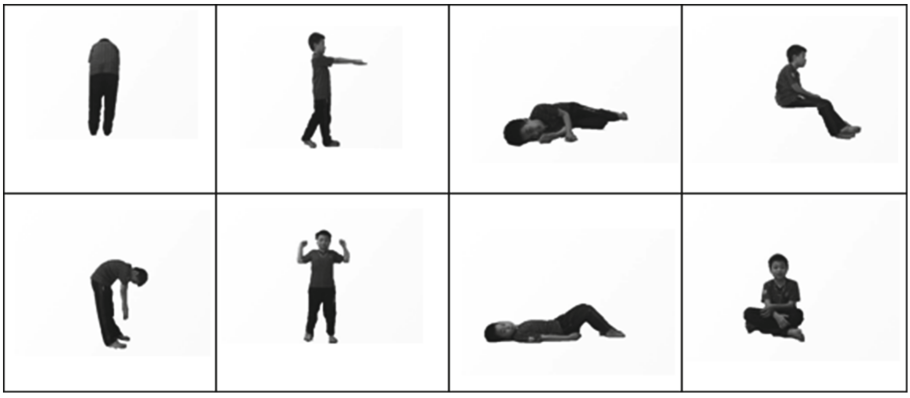


Fig. 4. Examples of the testing set

The testing set obtains from the training set by transforming images. Each posture includes 170 images. Some testing images are presented in Fig. 4.

4.2 Experimental Results

We conduct experiments on five classifiers including a non-linear SVM (our choice) [4], a Naïve Bayes classifier [13], Gaussian Mixture Model [6], Fuzzy Adaptive Resonance Theory (Fuzzy ART) [10], and K-mean [5]. These classifiers use two proposed features for classifying. A SVM optimizes the generalization error by the training process. A Naive Bayes classifier performs a decision based on probabilistic data. Gaussian Mixture Model makes statistical inferences about the properties of the sub-populations given only observations on the pooled population, without sub-population identity information. Fuzzy ART provides a

result based on the similarity of inputs and a threshold for deciding a category. K-mean also uses the similarity of inputs but excepting a threshold.

A confusion matrix is used to show the classification accuracy of categories. The confusion matrix is derived from a comparison of reference postures to the classified postures. This matrix takes the form of the columns representing the reference data by category and rows representing the classification by category.

We use a non-linear SVM in the one-against-all mode. The kernel is Gaussian radial basis function. We uses SVM functions of Matlab to perform sub-tests for each category. Training data of each test are shown in Table 2 and results are summarized in Table 3.

Table 2. The training dataset for each test of SVM

Test for	Bending	Sitting	Lying	Standing
Bending	10	3	3	3
Sitting	5	15	5	5
Standing	7	7	20	7
Lying	5	5	5	15

Table 3. The classifying accuracy of SVM

Test for	Number of correct patterns	Recognition rate
Test 1: Bending	160 + 426	94.1 %
Test 2: Sitting	160 + 296	94.1 %
Test 3: Standing	170 + 425	100.0 %
Test 4: Lying	165 + 494	97.1 %

Tables 4, 5, 6 and 7 show the accuracy of other classifiers. Table 8 illustrates the recognition rate of each method.

Table 4. The classifying accuracy of Naive Bayes classifier

		Reference data				Recognition rate
		Bending	Sitting	Standing	Lying	
Classified data	Bending	135	1	34	0	79.4 %
	Sitting	1	135	31	3	79.4 %
	Standing	5	4	161	0	94.7 %
	Lying	0	24	0	146	85.9 %

Table 5. The classifying accuracy of K-mean

		Reference data				Recognition rate
		Bending	Sitting	Standing	Lying	
Classified data	Bending	126	2	42	0	74.1 %
	Sitting	60	70	40	0	41.2 %
	Standing	0	0	170	0	100.0 %
	Lying	0	41	0	129	75.9 %

Table 6. The classifying accuracy Gaussian Mixture Model

		Reference data				Recognition rate
		Bending	Sitting	Standing	Lying	
Classified data	Bending	169	1	0	0	99.4 %
	Sitting	29	97	0	44	57.1 %
	Standing	139	31	0	0	0.0 %
	Lying	0	0	67	103	60.6 %

Table 7. The classifying accuracy of Modified Fuzzy ART

		Reference data				Recognition rate
		Bending	Sitting	Standing	Lying	
Classified data	Bending	96	29	0	45	56.5 %
	Sitting	0	92	78	0	54.1 %
	Standing	31	33	80	26	47.1 %
	Lying	46	41	41	42	24.7 %

Table 8. Classification rates of human body postures

Classifier	Recognition rate
SVM	96.3 %
Bayes classifier	84.9 %
K-mean	72.8 %
Gaussian Mixture Model	54.3 %
Modified Fuzzy ART	45.6 %

Data from Table 3 show the SVM classifies effectively for all postures with the accuracy from 94.12 % to 100 %. Table 8 shows the SVM is the best classifier.

5 Conclusion

We proposed two simple features to classify human body postures. These features are formulated from the square of human body. Then, we choose a non-linear SVM as a classifier to improve the ability of classifying. Experiment results show the chosen SVM classifies better than other methods (Naive Bayes classifier, Fuzzy ART, Gaussian Mixture Model and K-mean).

References

1. Juang, C.-F., Chen, T.-C., Du, W.-C.: Human body 3D posture estimation using significant points and two cameras. Sci. World J. **2014**, 17 (2014)
2. Juang, C.-F., Liang, C.-W., Lee, C.-L., Chung, I.-F.: Vision-based human body posture recognition using support vector machines. In: International Conference on Awareness Science and Technology, pp. 150–155 (2012)
3. Wu, C., Aghajan, H.: Model-based human posture estimation for gesture analysis in an opportunistic fusion smart camera network. In: IEEE International Conference on Advanced Video and Signal Based Surveillance, pp. 453–458 (2007)
4. Cortes, C., Vapnik, V.: Support vector networks. Mach. Learn. **20**(3), 273–297 (1995)
5. MacKay, D.: Information Theory, Inference and Learning Algorithms, pp. 284–292. Cambridge University Press, Cambridge (2003). ISBN: 0-521-64298-1
6. McLachlan, G., Peel, D.: Finite Mixture Models. Wiley, Hoboken (2000)
7. Cohen, I., Li, H.: Inference of human postures by classification of 3D human body shape. In: IEEE International Workshop on Analysis and Modeling of Faces and Gestures, pp. 74–81 (2003)
8. Shotton, J., Sharp, T., Kipman, A., Fitzgibbon, A., Finocchio, M., Blake, A., Cook, M., Moore, R.: Real-time human pose recognition in parts from single depth images. ACM Commun. **56**, 116–124 (2013)
9. Goldmann, L., Karaman, M., Sikor, T.: Human body posture recognition using MPEG-7 descriptors. In: Visual Communications and Image Processing, pp. 18–22 (2004)
10. Hoa, N.T., Bui, T.D.: A new effective learning rule of Fuzzy ART. In: Conference on Technologies and Applications of Artificial Intelligence, pp. 224–231. IEEE Press (2012)
11. Zerrouki, N., Houacine, A.: Automatic classification of human body postures based on the truncated SVD. J. Adv. Comput. Netw. **2**, 58–62 (2014)
12. Cucchiara, R., Grana, C., Prati, A., Vezzani, R.: Probabilistic posture classification for human-behavior analysis. IEEE Trans. Syst. Man Cybern. **35**, 42–54 (2005)
13. Russell, S., Norvig, P.: Artificial Intelligence: A Modern Approach, 2nd edn. Prentice Hall, Upper Saddle River (2003)
14. Girondel, V., Bonnaud, L., Caplier, A., Rombaut, M.: Belief theory-based classifiers comparison for static human body postures recognition in video. World Acad. Sci. Eng. Technol. **4**, 627–630 (2007)

Cluster Analysis, Classification and Forecasting Tool on DS30 for Better Investment Decision

Arshiful Islam Shadman, Sheikh Shadab Towqir, M. Ahnaf Akif, Mehrab Imtiaz, and Rashedur M. Rahman(✉)

Department of Electrical and Computer Engineering, North South University, Plot-15, Block-B, Bashundhara, Dhaka 1229, Bangladesh
{arshiful.islam,rashedur.rahman}@northsouth.edu, shadabto@gmail.com, uniahnaf@gmail.com, mehrab.i009@gmail.com

Abstract. An important aspect of finance is forecasting of stock returns. Relation can be established using fundamental information that is publicly available to predict future stock returns. This helps to extract knowledge from existing data and use them to make decisions by the investors. However other than only predicting future stock returns, our study can group stocks, choose stocks from the group to make a portfolio and make a better investment decision.

Keywords: Cluster analysis · Classification · Decision tree · Weka · Data mining techniques · Dhaka Stock Exchange · DS30 · Stock sectors · Investors

1 Introduction

Stock investors of Dhaka Stock Exchange and other international stock markets often use several analyzing and technical tools in which historical data related to stocks are fed and these tools give investors an estimation of the upcoming stock price. Sometimes public investors dive into the market without gaining proper knowledge. How often do these investors go through analyzing their trades by themselves before actually placing the request for trade? Data mining, also known as knowledge discovery, is the process of analyzing data to extract useful patterns and information. In this paper we will discuss and demonstrate data mining techniques like cluster analysis, classification using decision tree and open source forecasting tools in programs like Weka can be used to make better stock investment decisions.

2 Gathering and Cleaning Data

2.1 Data Source

First, we started looking for historical data of the Dhaka Stock Exchange. Financial websites like http://www.dsebd.org, http://www.stockbangladesh.com and http://www.lankabd.com were the main source of information. Some data were readily downloadable in Excel (*.xlsx) or Comma Separated Value (*.csv) format.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_22

	A	B	C	D	E	F	G	H	I	J	K	L
1	Company	Company Code	Sector	Annual EPS	P/B Ratio	P/E ratio (TO BE CALCULATED)	Dividend Yeild	Book value per share	Market Capital	Shareholder's Equity	ROA(return on assets)	ROE(return on equity)
2	City Bank	CITYBANK	Bank	3.51	1.06	6.25	6.94%	26.74	19,179,976,878.90	18,051,934,076	0.64%	4.42%
3	Islami Bank	ISLAMIBANK	Bank	3.25	1.04	7.65	5.93%	28.93	41,376,760,167.60	39,661,305,476	1.29%	16.70%
4	Pubali Bank	PUBALIBANK	Bank	3.34	0	5.06	7.06%	25.14	15,142,429,583.60	17,757,786,189	0.92%	9%
5	United Commercial Bank	UCB	Bank	4	0.98	4.28	11.98%	26.93	17,814,808,069.50	18,166,882,999	0.84%	9.29%
6	Brac Bank Ltd	BRACBANK	Bank	3.07	2.8	14.02	4.59%	27.19	30,924,927,195.60	11,025,261,642	0.44%	6.61%
7	IDLC Finanace ltd	IDLC	Financial Institut	5.81	2.08	9.14	4.42%	32.46	13,548,691,379.30	6,527,829,412	2.28%	20.95%
8	Aftab Automobiles	AFTABAUTO	Engineering	3.04	0.94	17.11	3.14%	54.1	4,911,071,248.60	5,198,108,811	2.54%	3.93%
9	BSRM Steels Ltd	BSRMSTEEL	Engineering	6.14	3.7	15.03	3.17%	25.66	32,468,625,000.00	8,768,853,872	4.12%	15.31%
10	Olympic Industries Ltd	OLYMPIC	Food and Allied F	7.96	16.83	38.91	1.01%	21.31	56,915,936,613.20	3,381,654,099	20.24%	37.07%

Fig. 1. DS30 overview

24	Date	Open	High	Low	Close	Volume	Typical Price	Distribution	Accumulated RSI	
25	18-November-15	540.7	540.7	527	529.5	117510	532.4333333	-76035.8824	-485658.206	40.14083
26	19-November-15	532	536.6	529	530.5	60978	532.0333333	-36907.7368	-522565.943	40.83334
27	22-November-15	531	534.6	525	526.5	128318	528.6333333	-83799.5102	-606365.453	38.89497
28	23-November-15	530	534.7	523	525.6	158526	527.7666667	-88070	-694435.453	38.45267
29	24-November-15	532	532	522	524.3	46454	526.2	-27297.7113	-721733.164	37.7843
30	25-November-15	524.5	528.9	518	520.2	65054	522.3666667	-38793.6697	-760526.834	35.678

Fig. 2. Company code ACI

2.2 Stock Data

Dhaka Stock Exchange contains 566 companies and 22 different sectors. Analyzing historical data of all these companies would be difficult. Hence we chose thirty companies enlisted in the DS30 index of Dhaka Stock Exchange. These thirty companies have market capitalization of over 500 million taka as of re balancing reference date. S&P Dow Jones Indices designed and developed the DS30 Index with effect from January 28, 2013. A maximum of six months of historical data was available for each of these thirty companies. Attributes such as Open, High, Low, Close and Volume were downloaded for each of these companies. Other information like company trading code, annual earnings per share, book value per share, P/B, P/E, Dividend Yield, ROE and ROA and sector information were also noted.

2.3 Attributes of Data

For fast data calculation and preparation Microsoft Office Excel was used. Data such as RSI, MACD, ADX, Bollinger Bands and Stochastic Indicators, Typical Price were calculated step by step as per their respective formulas. Those are described in Methodology section.

2.4 Data Cleaning

Due to unavailability of some data of the depended variables some instances were removed from consideration. In addition, for some companies there were no trades on some dates. There is a mismatch of trading dates within individual companies. Such dates were excluded from all thirty companies and only the common dates were considered for data analysis. Below we have provided sample snapshot of the data being used for analysis. Figure 1 shows the overview information of all the thirty companies we have chosen and Fig. 2 shows the data being downloaded and calculated for a company with trading code ACI.

3 Methodology

In this research, we propose a mechanism of using a combination of data mining techniques which will allow investors to make better investment decisions. Typically what most learned investors do is use technical financial tools. But here we will use clustering, decision tree and Weka's forecasting tool. We will apply all these techniques step by step with attributes of the data we have both downloaded and calculated. But before we dig deep let us look at what each of these techniques are in general:

3.1 Cluster Analysis

Clustering is a data mining technique that partitions a set of data into similar subclasses or clusters. Objects having similar characteristics are grouped together in the same cluster. We will use WEKA [5,6] for clustering. The necessity of clustering is to find the best group of companies based upon certain investors criteria. The stock market is full of several indicators and ratios that help investors choose companies. We will be using some of those ratios as the clustering attributes to find the cluster of companies that matches exactly with the threshold of these ratios. However, if we cannot find the best cluster we will choose the closest cluster based on the centroid values of the cluster.

3.2 Classification Using Decision Tree

Classification refers to examining the attributes of a new data set and assigning it to a target category or class. It uses training data to build a model which can be applied on unclassified data. We will use WEKA classify option for this research purpose. Stock market websites (mentioned previously) contains historical data. These historical data can be used to calculate several price indicators. These indicators can be used as attributes along with a class attribute for classifying whether a stock price will rise or fall.

3.3 Weka Forecasting Tool

Weka's forecasting tool is an effective tool which can forecast the future values of a data set. This is done by importing a time series forecasting model into Weka's forecasting environment. Weka's forecasting tool is not the same as the classification or regression techniques. Forecasting tool is essential because it is impossible for an investor to foresee the price of a stock tomorrow or in the near future.

4 Data Attributes

P/E Ratio (Price to Earning): P/E ratio is used to decide whether to buy shares of a company or not. The price-earnings ratio can be calculated as:

$$= \frac{currentPriceOfTheStock}{company'sEarningsPerShare}$$

P/B Ratio (Price to Book): P/B ratio is used to compare the market price of the share with a company's book value. The P/B ratio is calculated as:

$$= \frac{StockPrice}{TotalAssets - IntangibleAssetsandliabilities}$$

Dividend Yield: The dividend yield is the percentage of how much money a shareholder can make profit annually for buying the share. Formula for Dividend Yield:

$$= \frac{annualDividendPerShare}{stockPricePerShare}$$

ROA: It is basically how much money a company earns for every dollar in assets. The formula for return on assets is:

$$= \frac{netIncome}{totalAssets}$$

ROE: An indicator to measure whether the company is growing at an acceptable rate or not.

$$returnOnEquity = \frac{netIncome}{shareholder's Equity}$$

Open: It is the price at which the first trade is made by a security on a trading day.

High: The maximum price at which a stock is traded in the timespan of the trading day. It is usually higher than the opening or closing price.

Low: The minimum price at which a stock is traded in the timespan of the trading day. It is usually lower than the opening or closing price.

Close: The price at which the last trade is made by a security at the end of a trading session.

Volume: Volume is the aggregate of the number of shares that have been traded on a given trading day.

RSI: Using this indicator the strength and weakness of the stock is measured and it figures out whether a stock is overbought or oversold by which investor knows the gains or losses for a period of time.

$$RSI = 100 - \frac{100}{1 + RS}, \quad RS = Averageofxdays'upcloses/Averageofxdays'downcloses.$$

MACD: It uses the difference between long term and short term price trends to predict future price movements in the stock.

$$MACD = EMA12 - EMA26, \qquad EMA = ExponentialMovingAverage$$

Bollinger Band: This method is used by traders to detect the utmost short-term prices in a stock.

If prices are higher than the upper band = Overbought.
If prices are lower than the lower band = Oversold.

Stochastic Oscillator: It is a technical investment analysis tool that is used to measure a stock's closing price in comparison to its price range over a given period of time. The stochastic oscillator is determined by two numbers:

$$\%K = 100 * \frac{C - LofN}{HofN - LofN}, \qquad \%D = 3periodMovingAverageOf\%K$$

Average Directional Index (ADX): ADX is used in technical analysis to measure market trendlines. $ADX, DX+$ and $DX-$ are the three curves used to analyze ADX. If the curve of ADX is below 20 then it means that the market is not trending, and if it is above 20, it means trend may be starting.

5 Clustering

Investors often have to decide on which companies to invest. To do the so they build a portfolio and also see the profits returning from them. Cluster analysis of the companies with a chosen combination of attributes such as P/B, P/E, ROA, ROE etc. will provide the investor with clusters of companies and hence choose cluster(s) of companies for further analysis. Investor(s) can either select all the specified attributes for clustering (multi-dimensional data) or a combination of the subset of attributes. There are certain thresholds for each of these attributes which specifies the performance of a company. For simplicity, we will consider clustering on the following attributes:

1. Return on equity (ROE)
2. Return on assets (ROA)
3. Dividend Yield
4. Price-Earnings ratio (P/E)
5. Price-to-Book ratio (P/B)

ROE indicates if a company value is growing at an acceptable rate whereas ROA determines how much a company is earning per dollar of assets. According to [1] many investors and stock managers look for companies with ROE at least 15 % and ROA below 5 %. According to [2,3] a dividend yield between 5 % to 15 % and P/B under 3.0 is good for investment. Even though there is no such predefined

range or threshold for P/E ratio usage in stock selection but according to [4] company with P/E greater than 20 gives good returns. The Simple K-means Clustering (K = 5), using Euclidean distance as the distance measure, is done on twenty four of the thirty DS30 companies.

Due to data unavailability of information of six companies, they were not considered for clustering. Using the above thresholds for each attribute, the best cluster of companies is chosen for further analysis using classification. We copied the cluster centroids of all the five clusters and evaluated the value of the centroid with the investor's requirements (i.e. stock ratio thresholds). Each cluster is provided a score of one when its centroid value for the ratio being evaluated meets the threshold.

Observation:
Clusters that meet P/B threshold: 1, 2, 3, 4
Clusters that meet P/E threshold: 2
Clusters that meet Dividend Yield threshold: 1
Clusters that meet ROA threshold: 1, 2, 4
Clusters that meet ROE threshold: 0, 3

After all the clusters are scored, we can observe that there is no cluster with a score of 5. Therefore, no single cluster met all the requirements. However, there was a tie between two clusters (cluster-1, 2) as the centroids of these clusters were close to the stock ratio thresholds (Table 1).

Companies that fall within these two clusters were highlighted in green. According to the K-Means clustering process companies under cluster-1 and 2, amongst DS30, are good for investment (Fig. 3).

Table 1. Cluster center values for different attributes.

Cluster	0	1	2	3	4
P/B	13.7967	0.858	2.185	1.93	1.282
P/E	7.21	1.89	24.585	4.95	4.419
DY	0.0182	0.0869	0.0301	0.0293	0.0304
ROA	0.2091	0.0198	0.0476	0.1656	0.041
ROE	0.5404	0.113	0.0714	0.2671	0.1008
SCORE	1	3	3	2	2

	A	B	C	D	E	F
1	Company Code	P/B Ratio	P/E ratio	Dividend	ROA	ROE
2	CITYBANK	0.99	2.66	0.06881	0.0064	0.0442
3	ISLAMIBANK	0.95	2.36	0.06438	0.0129	0.167
4	PUBALIBANK	0	1.2	0.07059	0.0092	0.0942
5	UCB	0.95	0.93	0.12195	0.0084	0.0929
6	BRACBANK	0	3.55	0.04751	0.0044	0.0661
7	IDLC	2.24	2.31	0.04371	0.0228	0.2095
8	AFTABAUTO	0.91	5.97	0.03187	0.0254	0.0393

Fig. 3. DS30 attribute values

6 Classification Using Decision Trees

To further narrow down our results and decide whether to invest in any of the companies found from clustering, we apply classification using J48 algorithm (decision tree) in addition with the forecasting tool of Weka on the group of companies. For now we have chosen the following stock market indicators as the main attributes for the construction of the decision tree:

1. Relative Strength Index (RSI)
2. Moving Average Convergence Divergence (MACD)
3. Average Directional Index (ADX)
4. Bollinger Bands (Upper & Lower Bollinger Bands)
5. Stochastic Indicator

We choose all the above mentioned attributes to generate our model. All the above attributes we calculated in Microsoft Office Excel step by step in accordance to their respective formulas. "RISE/FALL" is considered to be the class attribute. To determine the nominal value of the class attribute (RISE or FALL) we used the Typical price of a single day where,

$$typicalPrice = \frac{High + Low + Close}{3}$$

If the typical price of the current day is greater than the previous day then the class attribute is set to RISE, else the class attribute is set to FALL. The following Excel IF code is used here:

$$Fx = IF(G5 > G4, \text{“}RISE\text{”}, \text{“}FALL\text{”})$$

G5: Today's Typical Price
G4: Previous Day's Typical Price

All the deciding attributes and the class attributes were copied to a new Excel file and exported as CSV (Comma Delimited Value) file so that it could be used in Weka. The following snapshot gives an idea of the training data being used as the classification purpose. Figure 4 shows only four records out of ninety day data of CITYBANK. The decision tree in Fig. 5 is obtained when all the ninety day data was being fed into Weka Classification tool. When all the ninety day record of CITYBANK is considered, the model (Fig. 5) generated can classify 84/90 instance correctly. This displays an accuracy of 93.3 %. To further analyze

	A	B	C	D	E	F	G
1	RSI	MACD	ADX	UPPER BOLLINGER BAND	LOWER BOLLINGER BAND	STOCHASTIC	RISE/FALL
2	44.6352	-0.03853	21.3151	20.79729915	19.27270085	48.71794872	FALL
3	44.6352	-0.05247	20.04029	20.77005319	19.35994681	42.10526316	FALL
4	42.46237	-0.06645	19.25949	20.7315606	19.4484394	28.57142857	FALL
5	42.46237	-0.07319	18.07301	20.69957612	19.52042388	18.75	RISE

Fig. 4. Different attribute values

we decided to split the ninety day record into 80:10 where eighty day record is considered to be the training data set and the rest as the test data set. Figure 6 shows the decision tree obtained considering the eighty day record as the training data set. The classifier output obtained after splitting the ninety day record into 80:10 shows that 7/10 record are correctly classified, providing an accuracy of seventy percent with a root mean squared value of 0.4956. Figures 5 and 6 are both decision trees of CITYBANK. For comparison purpose we took two other companies into consideration and compared their classifier accuracy. Table 2 depicts the accuracy.

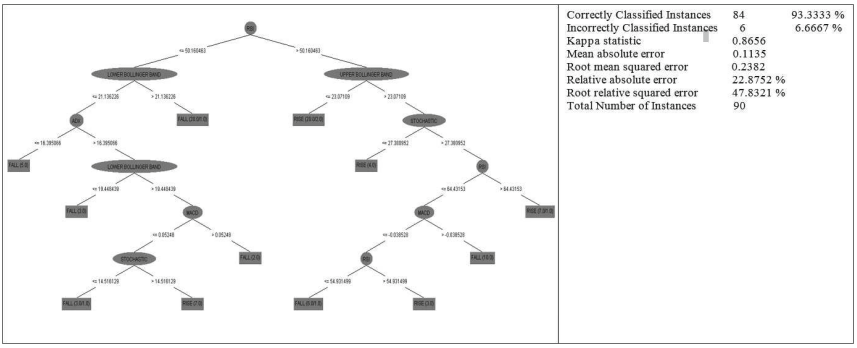


Fig. 5. Tree obtained when all records are used for training and summary of the run information

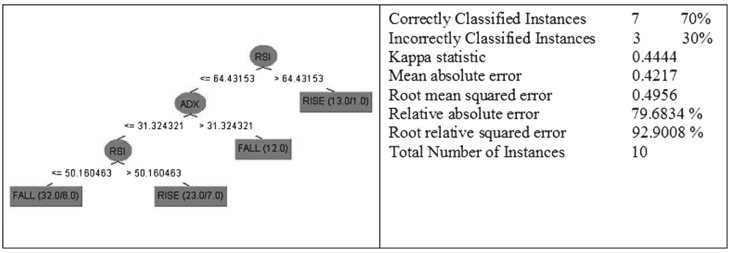


Fig. 6. Tree when 80:10 distribution for training and testing, and summary of the run information

Table 2. Accuracy of classification

Company code	% of correctly classified instances (without splitting)	% of correctly classified instances (with 80:10 splitting)
CITYBANK	93.33 %	70 %
BSCCL	87.87 %	50 %
ISLAMIBANK	85.56 %	70 %

7 Integration of Weka's Forecasting Tool and Classification Using Decision Tree

As a trader, one would like to know what the market price and the corresponding price indicator of a stock will be tomorrow or in the near future? Will there be an upward trend or downward trend? Will the price rise or fall? It is not possible to know where the market will stand. The stock market is highly unpredictable. Several economic factors such as interest rates, inflation, deflation, economic and political shocks and changes in economic policies etc. can cause the price of a stock to rise or fall. However, tools like Weka's forecasting tool can give an estimation of the market price and price indicators. The objective of this particular section is to use the historical price of three companies and use them to forecast. The forecasted value will be treated as the test dataset for the model obtained from the historical dataset and predict whether the price will rise or fall. We use SMOReg for prediction which is SVM (Support Vector Machine) implementation for regression. More details about it could be found elsewhere [7].

In order to complete this section we took eighty-nine day record to predict the value of RSI, MACD, ADX, UPPER BOLLINGER BAND, LOWER BOLLINGER BAND and STOCHASTIC INDEX for the 90th day. We already have the actual data for the 90th day.

The information below shows the actual and predicted data for the 90th day of CITYBANK, ISLAMIBANK and BSCCL:

Company	RSI	MACD	ADX	UPPER BOLLINGER BAND	LOWER BOLLINGER BAND	SLOW STOCHASTIC	RMSE
BSCCL(actual)	45.096663	-0.1209697	10.893294	108.1858968	102.9441032	19.60784314	
BSCCL(predicted)	46.2391	-0.1857	9.6905	109.0321	102.4348	17.4906	1.170059
CITYBANK(actual)	49.086699	0.05328785	17.77276906	22.72036878	20.88963122	69.56521739	
CITYBANK(predicted)	50.1893	0.0636	19.0623	22.676	21.0724	67.9461	0.960524
ISLAMIBANK(actual)	55.166773	0.3386993	30.4661521	26.26272574	21.41727426	58.64197531	
ISLAMIBANK(predicted)	55.7762	0.4167	35.2347	26.4433	21.1791	67.2449	4.025257

Fig. 7. RMSE between actual and predicted data

The figure above shows the root mean square value of three companies for the forecasted value and the actual values of the indicators of the 90th day. This provides a measure for the prediction accuracy of Weka's Forecasting tool. Since lower values of RMSE indicate better fit, we can say that the predicted value of CITYBANK is the closest to actual value relative to the other two companies (Fig. 7).

8 Conclusion

This paper is an integration of cluster analysis, classification using decision tree and using Weka's forecasting tool. Simple K-Means clustering technique was applied on thirty companies of the DS30 of the Dhaka Stock Exchange on to find the best cluster of companies based on the threshold of the clustering attributes. Classification using J48 algorithm to generate decision tree (model) is applied

on one of the selected group of companies. This gave us results of both correctly and incorrectly classified instances of one company using the model generated. Later, Weka's forecasting tool was used to predict near future values of the price indicators.

References

1. McClure, B.: ROA and ROE Give Clear Picture Of Corporate Health. http://www.investopedia.com/articles/basics/05/052005.asp?o=40186&l=dir&qsrc=1&qo=serpSearchTopBox&ap=investopedia.com
2. McClure, B.: Using Price to Book Ratio to Evaluate Companies. http://www.investopedia.com/articles/fundamental/03/112603.asp?o=40186&l=dir&qsrc=999&qo=investopediaSiteSearch
3. Moskowitz, D.: Top Dividend Stocks: How To Pick Them. http://www.investopedia.com/articles/active-trading/042315/top-dividend-stocks-how-pick-them.asp
4. Kalarani, C., Indrakala, S., Nachiaar, K.: Analysis of nifty fifty stock based on K-Means clustering technique for stock market prediction. Int. Res. J. Math. Eng. & IT **2**(7), 26–33 (2014)
5. Kumar, R., Blara, A.: Time Series Forecasting of Nifty Stock Market Using Weka. http://jrps.in/uploads/March%202014/research_paper_raj.pdf. Accessed 20 Oct 2016
6. WEKA Data Mining Tool. http://www.cs.waikato.ac.nz/ml/weka/
7. SMOReg. http://weka.sourceforge.net/doc.dev/weka/classifiers/functions/SMOreg.html

Comparing Modified PSO Algorithms for MRS in Unknown Environment Exploration

Anh-Quy Hoang(✉) and Minh-Trien Pham(✉)

VNU University of Engineering and Technology,
144 Xuan Thuy, Cau Giay, Hanoi, Vietnam
{quyha,trienpm}@vnu.edu.vn

Abstract. Multi-robot systems (MRS) have shown clear advantages over single robots in the application of exploring unknown environments - a fundamental problem in robotics. Among algorithms which are able to be applied to MRS in the application, Particle Swarm Optimization (PSO) - a heuristic optimization technique inspired by social behavior of natural swarms - has received much attention and is well-known for its efficiency and simplicity to implement. However, when conventional PSO is applied, the problems of disconnection and collision within the system are inevitable. Two of various methods proposed to address these crucial issues are applying BOIDS and Artificial Potential Field (APF) to modify PSO. In this work, we simulated both modified algorithms on Matlab under various scenarios for analysis and comparison.

1 Introduction

1.1 Multi-robot Systems

A Multi-robot System (MRS) is a set of robots operating in the same environment as an entity [1]. MRSs are applied in a wide range of applications such as military services, rescue operation, monitoring, surveillance, etc. [2] and especially environment exploration. Thanks to the ability to share information to cooperate and coordinate, MRSs are not only much more robust to local optima [3] but also able to achieve a more widespread coverage and a higher degree of accuracy in comparison with a single robot in exploration tasks.

1.2 Particle Swarm Optimization Algorithm

Since its first introduction in 1995 by Russel Eberhart and James Kennedy [4], PSO has continuously grown in popularity among bio-inspired heuristic algorithms. It is proved to be both simple to implement and efficient in optimization problems. PSO is now applied in various fields in engineering and technology [3].

PSO is particularly effective in finding local and global optima. In the searching process, the swarm disperse randomly over the search space before exploring. During the exploration, each particle is controlled by a velocity comprised of three components, namely inertial velocity, cognitive velocity, and social velocity [4].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_23

The three components keep particles' movement smooth while guides each particle to the best position of its own and of the whole swarm. Social learning factor should be increased and cognitive factor should be decreased throughout the searching process to maximize the exploratory capability of the entire swarm. Four stages in PSO are described below:

1. Initializing: To generate the population and evaluate the objective (fitness) function.
2. Updating personal best and global best: To check each particle, compare current fitness with best fitness for new personal best and global best.
3. Updating velocity and position of all particles: The position and velocity of each particle are updated using the following equations:
   New velocity = inertial velocity + cognitive velocity + social velocity

$$\mathbf{v}_{PSOi}(t) = w\mathbf{v}_i(t-1) + a_1 u_d(\mathbf{p}_i(t-1) - \mathbf{x}_i(t-1)) + a_2 U_d(\mathbf{g}(t-1) - \mathbf{x}_i(t-1)) \tag{1}$$

   New position = current position + new velocity
   (Note that in PSO both position and velocity are vectors.)

$$\mathbf{x}_i(t) = \mathbf{x}_i(t-1) + \mathbf{v}_i(t) \tag{2}$$

   where:
   w: Inertial factor, control the effect of the previous velocity on the new velocity.
   a_1, a_2: Cognitive and social learning factors.
   u_d, U_d: Uniform random number in [0,1].
4. Checking termination criteria: The searching process is terminated if:
   (i) Current step is equal to the predefined maximum step.
   (ii) The swarm has converged (Radius of the swarm is smaller than 10^{-3} % size of the search space).
   Otherwise, we come back to step 2 [4].

   Convergence rate of PSO is proved to be high. The apparent advantages of PSO over other heuristic algorithms make it a good choice for MRSs in the exploration mission. However, disconnection and collision within the system are unavoidable when PSO is applied for MRS. To address these problems, a large number of methods have been proposed. In this paper, we compare two of them: BOIDS and APF, in term of effectiveness in finding targets and capability of avoiding obstacles and remaining connectivity.

1.3 BOIDS

The term "BOIDS" is short for "bird-oid object" - bird-like object. BOIDS is an artificial life program simulating swarm behavior of birds, in which the entire swarm move relatively fast without collision. BOIDS has been applied widely in computer animations and motion pictures. It was first introduced by Crag Reynolds in 1986 and appeared in the proceedings of the ACM SIGGRAPH conference in 1987 [5]. BOIDS model is executed with a set of three simple rules.

The rules are used to steer the particles so that they could avoid collision, move in the same direction as their neighbors and maintain formation of the swarm.

The velocity of each particle is the sum of three velocity components: separation, alignment, and cohesion velocity. Each of them serves a purpose.

Separation velocity assures that a particle does not collide with any other particles. There is a separation region around each particle, the radius of this region is called "separation radius". If another particle invades this region, there will be a repulsive force between the two particles. The force results in a velocity vector with a magnitude corresponding to the depth of invasion.

Alignment velocity steers the velocity of a particle toward those of its neighbors, so that the swarm does not disperse while moving. It is obvious that each particle has a unique velocity, but in reality, it seems that the entire swarm moves in only a direction. This effect is created by alignment velocity. This component velocity is calculated by taking the average velocity of all neighboring flock mates.

Cohesion velocity preserves the formation of the swarm during the searching process. In every step, each particle is affected by a velocity component directing it into the center of its neighbors (average position or center of mass). Thus, the swarm keeps both its formation and flexibility [5].

1.4 Artificial Potential Field

Proposed by Oussama Khatib in 1986 for single robot path planning [6], APF is now widely used in works on MRS. Dissimilar to methods based on B-Splines [7] or behavior classification [8], APF generates around each robot a virtual potential field containing a repulsive field and an attractive field. The attractive field directs each robot towards other robots in the system to remain system connectivity while the repulsive field keeps them from colliding with other robots or obstacles. The combination of these two fields forms a control rule for MRS. There are multiple architectures for APF, in accordance with requirements of specific tasks. The magnitude of potential forces exerted on each robot are continuously updated based on the information it gets from the immediate surrounding environment, or from other robots via connection network. This model is inspired by Artificial Physics with quadratic functions. Therefore, artificial potential forces are used to remain the relation between robots in term of position. Potential forces are categorized into two main groups: passive forces and active forces. Passive forces are generated when robots emit signal and determine distances to neighboring robots or obstacles by the magnitude of reflected signal to avoid obstacles or remain relative position with other robots. The signal used in the application could be infrared, ultrasound, laser or camera. On the contrary, active forces are from outside sources, usually by other robots and transmitted via communication system. We use passive forces in the simulation of APF-PSO in this paper.

2 Methodology

Applying BOIDS and APF to MRS is not a new idea. However, in previous works, they are only used for formation control and path planning [9–12]. We are among the first groups to modified PSO by integrating these two methods into PSO for the task of space exploration [13–15]. In this paper, we do the simulation of BOIDS-PSO and APF-PSO then compare the results.

2.1 BOIDS-PSO and APF-PSO

Control laws for the modified algorithms are presented as following:

BOIDS-PSO

BOIDS model is combined with PSO to form a modified PSO algorithm, which is able to cope with the problems of collision and disconnection within the swarm. BOIDS-PSO velocity is a combination of BOIDS velocity and PSO velocity, i.e. it consists of six velocity components, three from PSO algorithm and three from BOIDS model.

BOIDS-PSO control rule is described by the following equation:

$$\mathbf{v}_i(t) = \mathbf{v}_{PSOi}(t) + \mathbf{v}_{BOIDSi}(t) \tag{3}$$

$\mathbf{v}_{PSO}$ has been discussed previously. $\mathbf{v}_{BOIDS}$ is comprised of three components:

$$\mathbf{v}_{BOIDSi}(t) = \alpha \mathbf{v}_{Ai}(t) + \beta \mathbf{v}_{Ci}(t) + \gamma \mathbf{v}_{Si}(t) \tag{4}$$

where:
$\mathbf{v}_{Ai}(t)$: alignment velocity
$\mathbf{v}_{Ci}(t)$: coherence velocity
$\mathbf{v}_{Si}(t)$: separation velocity
α, β, γ: adjusting parameters

APF-PSO

PSO is used for the explorative purpose, and APF is used to remain connectivity and avoid collision. Therefore, the control law applied to the MRS is modified PSO, with the introduction of APF. As discussed in the last section, there are multiple forms for APF according to specific problems. For the task of exploring unknown environment using MRS, we utilize the following formula:

$$\mathbf{F}_{APF12} = \frac{Gm_1m_2\mathbf{r}_{12}}{r_{12}^3}[u(r_{12}) - u(r_{12} - r_1) - ku(r_{12} - r_2) + ku(r_{12} - r_3)] \tag{5}$$

Where G and k are predetermined constants to regulate the magnitude of potential forces, m_1 and m_2 are respectively the weight of robot 1 and robot 2, r_{12} is the distance vector from robot 1 to robot 2, $\mathbf{F}_{APF12}$ is the force exerted by

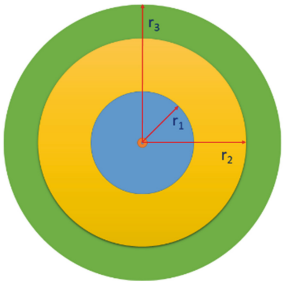


Fig. 1. Potential fields surrounding a robot

potential field of robot 1 on robot 2. The constants in Heaviside step functions, r_1, r_2, and r_3 are radii of the repulsive zone, the free zone, and the attractive zone respectively. r_3 must not larger than the communication range of a robot. The force field in each zone is a potential field, but it is not true to the entire surrounding space of a robot. Within the potential fields of a robot, other robots are thrust away if they are in the repulsive zone, pull towards if they are in the attractive zone; or receive no force if they are in the free zone (Fig. 1).

Total force exerted on robot j in the system is:

$$\mathbf{F}_{APFj} = \sum_{i=1}^{N} \mathbf{F}_{APFij} \tag{6}$$

Here, if $i = j, \mathbf{F}_{APFij} = 0$.

The force gives each robot an additional velocity component - APF velocity. The magnitude of this component is proportional to that of total APF force. Thus, in each step, the velocity of a robot is determined by:

$$\mathbf{v}_i(t) = \mathbf{v}_{PSOi}(t) + \mathbf{v}_{APFi}(t) \tag{7}$$

2.2 Comparing Modified PSO Algorithms

We simulate exploration process of the MRS in two scenarios and then make comparisons regarding collision probability, algebraic connectivity, and effectiveness when each algorithm is applied.

Scenario 1: In this scenario, the ability to avoid collision of MRS is evaluated. Robots and static obstacles are randomly distributed in search space in a manner that there exists no position overlap. Communication range is unlimited.

Scenario 2: We check whether the system is connected during the exploration. Robots are distributed such that the system is connected, i.e., the MRS's algebraic connectivity is initially positive. Communication range is limited.

2.3 Configuration of the MRS

In this study, we simulate a homogeneous MRS with the radius of each robot (r) being a unit of length. The system has direct communication and population size varies.

If we set r = 1, maximum velocity is 0.5/step. For APF-PSO, r_1 is 5, r_2 is 8 and r_3 is 12. For BOIDS-PSO, the communication range is 10 (scenario 2), the separation radius is 8.

3 Simulations and Results

3.1 Simulation Set-Up

Initial distribution of MRS in the search space is random (scenario 1) or predetermined (scenario 2). There are two static obstacles at (−35,35) and (−21, −31). Search space is limited by four lines: x = 50, x = −50, y = 50, y = −50. Objective functions are scaled Three-Hump Camel Function, Bohachevsky Function and Rosenbrock Function.

3.2 Comparison of Collision Avoidance Ability

The following tables show the probability of collision for both modified algorithms with different objective functions and number of robots in the swarm after 100 runs (i.e. 20000 steps) (Tables 1, 2 and 3).

Table 1. Collision probability - Bohachevsky function

	5 robots	10 robots	15 robots
BOIDS-PSO	1.1429e-05	3.2381e-05	1.0429e-04
APF-PSO	1.1667e-05	2.7059e-05	9.0909e-06

Table 2. Collision probability - Three-Hump Camel function

	5 robots	10 robots	15 robots
BOIDS-PSO	1.0952e-05	3.3810e-05	9.8571e-05
APF-PSO	3.1667e-05	3.4706e-05	1.7879e-05

It can be seen that the MRS's collision probability does not clearly differ between the two algorithms. When BOIDS-PSO is applied, with a small number of robots, the probability tends to be smaller, while with a larger number of robots (15), the probability is about ten times larger than the case when APF-PSO is applied.

Table 3. Collision probability - Rosenbrock function

	5 robots	10 robots	15 robots
BOIDS-PSO	6.1905e-06	4.4286e-05	9.8095e-05
APF-PSO	1.0000e-05	7.6471e-06	1.8182e-05

3.3 Comparison of Connectivity Maintenance Ability

With each modified algorithm, we run the simulation for 100 times and analyze algebraic connectivity (second smallest eigenvalue of Laplace matrix of the graph whose nodes are represented by robots in MRS). Gathered data are provided in the tables below (Tables 4, 5 and 6):

Table 4. Algebraic connectivity - Bohachevsky function

	Mean of mean algebraic connectivity	Standard deviation of mean algebraic connectivity
BOIDS-PSO	1.1425	0.0855
APF-PSO	6.7154	0.6087

Table 5. Algebraic connectivity - Three-Hump Camel function

	Mean of mean algebraic connectivity	Standard deviation of mean algebraic connectivity
BOIDS-PSO	0.8933	0.1608
APF-PSO	6.0456	0.5092

Table 6. Algebraic connectivity – Rosenbrock function

	Mean of mean algebraic connectivity	Standard deviation of mean algebraic connectivity
BOIDS-PSO	0.8997	0.1376
APF-PSO	8.2091	0.4610

For all objective functions, both the mean of mean algebraic connectivity and standard deviation of that quantity are smaller when BOIDS-PSO is applied, indicating that the MRS is not closely connected but the state of connection is more stable with BOIDS-PSO.

3.4 Comparison of Effectiveness

In this section, we run simulations to determine how effective each algorithm is. After each simulation, the swarm found a best value. We run simulation for 100 times and compare the mean of best value found using each algorithm. Besides, the best values found after 100 runs are given (the real minimum value is 0) (Tables 7, 8 and 9):

Table 7. Effectiveness – Bohachevsky function

	Best of best fitnesses	Mean of best fitnesses
BOIDS-PSO	0.0006	0.1468
APF-PSO	1.9857e-04	1.1400

Table 8. Effectiveness – Three-Hump Camel function

	Best of best fitnesses	Mean of best fitnesses
BOIDS-PSO	0.0001	0.0387
APF-PSO	1.8810e-05	0.2789

Table 9. Effectiveness – Rosenbrock function

	Best of best fitnesses	Mean of best fitnesses
BOIDS-PSO	0.0023	0.1932
APF-PSO	1.4660e-04	0.0567

The value of mean of best fitnesses is of the same scale for both algorithms, while the best value is about 10^4 times smaller when APF-PSO is applied for the MRS, that mean the overall effectiveness of BOIDS-PSO and APF-PSO are almost the same, but the variance of best values when we use APF-PSO are much higher. This means BOIDS-PSO gives a more stable accuracy, but APF-PSO can always claim higher accuracy in the condition that the algorithm is run for numerous times.

4 Conclusions and Future Works

In this paper, through simulations we have compared BOIDS-PSO and APF-PSO, two modified PSO algorithms used to address the problems of collision and disconnection of MRS during exploration missions. The results indicate that BOIDS-PSO and APF-PSO are of the same level of accuracy, though APF-PSO can be used to achieve higher accuracy in some cases. The collision probability

of MRS are similar in both cases: BOIDS-PSO or APF-PSO is applied. The connectivity of the system is more stable when BOIDS-PSO is utilized but stronger when APF-PSO is used.

Future studies will be conducted to evaluate the effectiveness of different algorithms for MRS and enhance the performance of MRS in exploration missions.

Acknowledgments. This work has been supported by Vietnam National University, Hanoi (VNU), under Project No. QG.15.25.

References

1. Farinelli, A., Iocchi, L., Nardi, D.: Multirobot systems: a classification focused on coordination. IEEE Trans. Syst. Man Cybern. Part B (Cybern.) **34**(5), 2015–2028 (2004)
2. Robin, M., Blitch, J., Casper, J.: AAAI/RoboCup-2001 Urban Search and Rescue Events. AI Magazine 23, no. 1: 37 (2002)
3. Couceiro, M.S., Rocha, R.P., Ferreira, N.M.: A novel multi-robot exploration approach based on particle swarm optimization algorithms. In: 2011 IEEE International Symposium on Safety, Security, and Rescue Robotics, pp. 327–332. IEEE, November 2011
4. Kennedy, J.: Particle swarm optimization. In: Encyclopedia of Machine Learning, pp. 760–766. Springer, US (2011)
5. Reynolds, C.W.: Flocks, herds and schools: a distributed behavioral model. ACM SIGGRAPH Comput. Graph. **21**(4), 25–34 (1987)
6. Khatib, O.: Real-time obstacle avoidance for manipulators and mobile robots. Int. J. Robot. Res. **5**(1), 90–98 (1986)
7. Paulos, E.: On-line collision avoidance for multiple robots using b-splines. University of California, Berkeley, Computer Science Division (1998)
8. Kumar, R.A., Menon, A.: Collision avoidance in a multi-robot system by emulating human behaviour. In: International Conference on Information Systems Analysis and Synthesis and World Multiconference on Systemics, June 2001
9. Sharma, S., Tiwari, R.: A survey on multi robots area exploration techniques and algorithms. In: 2016 International Conference on Computational Techniques in Information and Communication Technologies (ICCTICT), pp. 151–158. IEEE, March 2016
10. Balch, T., Hybinette, M.: Social potentials for scalable multi-robot formations. In: Proceedings of the IEEE International Conference on Robotics and Automation, ICRA 2000, vol. 1, pp. 73–80. IEEE (2000)
11. Guanghua, W., Deyi, L., Wenyan, G., Peng, J.: Study on formation control of multi-robot systems. In: 2013 Third International Conference on Intelligent System Design and Engineering Applications (ISDEA), pp. 1335–1339. IEEE, January 2013
12. Lee, L.F., Bhatt, R., Krovi, V.: Comparison of alternate methods for distributed motion planning of robot collectives within a potential field framework. In: Proceedings of the 2005 IEEE International Conference on Robotics and Automation, pp. 99–104. IEEE, April 2005

13. Hoang, A.Q.: Simulating Hybrid BOIDS-PSO Algorithm for MRS in Unknown Environment Exploration. (Unpublished bachelor thesis) VNU University of Engineering and Technology, Hanoi, Vietnam (2015)
14. Hoang, A.Q., Pham, M.T.: Light source detection using multirobot systems with particle swarm optimization approach. In: The 3rd Vietnam Conference on Automation and Control (2015)
15. Hoang, A.-Q., Pham, M.-T.: Swarm intelligence-based approach for macroscopic scale odor source localization using multi-robot system. In: Akagi, M., et al. (eds.) Advances in Information and Communication Technology. AISC, vol. 538, pp. 593–602. Springer, Switzerland (2016)

Design Adaptive-CTC Controller for Tracking Target Used Mobile Robot-Pan Tilt-Stereo Camera System

Le Van Chung(✉) and Duong Chinh Cuong

University of Information and Communication Technology,
Thai Nguyen University, Thai Nguyen, Vietnam
{lvchung,dccuong}@ictu.edu.vn

Abstract. The paper presents a dynamic model of mobile robot - pan tilt robot - stereo camera system. Stereo camera is placed on the head of pan tilt robot and carried by a mobile robot. The pan tilt robot will rotate to keep the image target on image frame and take the moving information of mobile robot out. The main contribute of this paper is designed an Adaptive Computed-Torque-Control controller for the dynamic system. The controller is used to control pan tilt robot track the moving target and mobile robot move to reach to the target when there are noise effects.

Keywords: Adaptive controller · Mobile robot · Pan tilt · Stereo camera · Tracking target

1 Introduction

Mobile robot – pan tilt robot – stereo camera system consists of is a Pioneer 3DX mobile robot and two degree-of-freedom pan tilt robot including stereo camera. This is a complex integrated system with many influences between links [1]. If we only have a mobile robot [2] or pan/tilt robot, we can only control mobile robot moving follow an anticipate path or track a moving target [3]. A combined system can track and move to the moving target.

Some authors usually construct separation the dynamic of mobile robot and dynamic of other element [11]. Recently, the dynamic of integrated system has been constructed [1] for the pan tilt robot with a camera placed on mobile robot. Nevertheless, operating space of robot is restricted because the target distance with mono camera is not measured. In this paper, stereo camera is used to obtain not only the image feature but also the distance to the target.

Control algorithms can be classified as classical kinematic controller [3,10]; advance dynamic controller [1,2]; intelligent control using neural network [1]. Optimization methods are used in [5,9] as Linear Quadratic Regulator (LQR) method [4,7] or integrated with Kalman filter as Linear Quadratic Gaussian (LQG) control [8] for the system impacted of noise.

The task of this paper is to keep the target image in the camera frame while the target moves in space with velocity vector changing and mobile robot move

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_24

to reach the target while the system has the effect of noise. The Liner Quadratic Regulator (LQR) controller is used to optimize the energy consumption and position error. The paper is organized in the following way. In Sect. 2, the kinematic control problem for their system is solved. In Sect. 3, a dynamic control algorithm with LQR controller and the asymptotic stability of the overall system is proved by Lyapunov stability method. Section 4, simulation results on Matlab-Simulink are showed to demonstrate the performance of the proposed control algorithm. Finally, in the Sect. 5, some conclusions are exposed.

2 Construct Model of Robot - Pan/Tilt - Stereo Camera System

2.1 Construct Kinematic Model

In Fig. 1, the coordinates are assigned to mobile robot, pan/tilt robot and cameras. We have the homogeneous transformation matrix which transforms coordinates from $O_cX_cY_cZ_c$ to $O_0X_0Y_0Z_0$ is:

$$ {}^0R_C = \begin{bmatrix} -s_{12} & -c_{12}s_3 & c_{12}c_3 & a \\ c_{12} & -s_{12}s_3 & s_{12}c_3 & b \\ 0 & c_3 & s_3 & c \\ 0 & 0 & 0 & 1 \end{bmatrix}, \tag{1} $$

$a = c_3(x_m + l_1c_1) + s_2s_3(y_m + l_1s_1) + l_2s_3; \quad b = c_2(y_m + l_1s_1) + l_3c_3;$
$c = -s_3(x_m + l_1c_1) + s_2c_3(y_m + l_1s_1) + l_2c_3 + l_3s_3; \; c_3 = \cos(\theta_3); \; s_3 = \sin(\theta_3)$
$c_{12} = \cos(\theta_1)\cos(\theta_2); \; c_2 = \cos(\theta_2); \; s_2 = \sin(\theta_2)$

Assumption 1: The motion of robot wheels is non-slip. Denotation the angular and straight velocity of the two wheels are (ϕ_{left},$\dot{\phi}_{left}$, v_{left}), (ϕ_{right},$\dot{\phi}_{right}$, v_{right}). The relationships between v_{left}, v_{right} and straight velocity v_r angular velocity ω_r of robot are:

$$ \mathbf{v} = \begin{bmatrix} v_{right} \\ v_{left} \end{bmatrix} = \frac{1}{r} \begin{bmatrix} 1 & k/2 \\ 1 & -k/2 \end{bmatrix} \begin{bmatrix} v_r \\ \omega_r \end{bmatrix}. \tag{2} $$

where $r = d/2$ is the radius of the wheel; k is the distance between two wheels. Because the movement of the robot is a plane-parallel motion, the relationship between the desired position and it's velocity v_{rd}, ω_{rd} is:

$$ x_{rd} = \int v_{rd}\cos\theta_{1d}; \;\; y_{rd} = \int v_{rd}\sin\theta_{1d}; \;\; \theta_{1d} = \int \omega_{rd}. \tag{3} $$

2.2 Construct Dynamic Model

We used Lagrange method to construct the dynamic equation of the system. Denote m_1, m_2, m_3, m_w respectively is a mass of mobile robot, joints pan, tilt joints with stereo camera and wheel. θ_1 is the direction angle of the mobile robot. Pan angle is θ_2, its rotation around the axis Z_1, Tilt angle is θ_3,

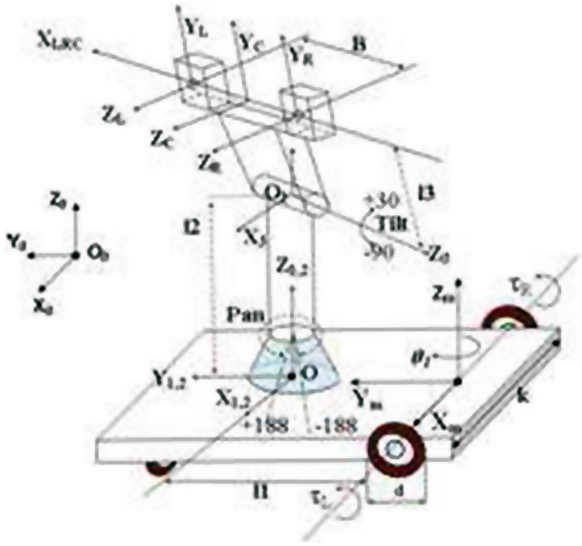


Fig. 1. Structure of mobile robot- pan/tilt-stereo camera

its rotation around the axis Z_3. is the current coordinate of mobile robot in O_0. The total kinetic and potential energy of the system showed in (4) and (5):

$$\begin{aligned} K &= 0.25\left[I_{3y}(1+\cos(2\theta_3)) + I_{3x}(1-\cos(2\theta_3)) + 2I_{2z}\right](\dot{\theta}_1+\dot{\theta}_2)^2 \\ &+ 0.5(m_1+m_2+m_3)(\dot{x}_r^2+\dot{y}_r^2) + 0.5\left[I_{1z}+2I_d+(m_2+m_3)l_1^2\right]\dot{\theta}_1^2 \\ &+ 0.5((I_{3z}+m_3l_3^2)\dot{\theta}_3^2 + (m_{\mathrm{w}}r^2+I_{\mathrm{w}})(\dot{\phi}_R^2+\dot{\phi}_L^2)) - m_3l_1l_3c_2s_3(\dot{\theta}_1\dot{\theta}_2+\dot{\theta}_3^2) \end{aligned} \tag{4}$$

$$P = m_3gl_3\sin(\theta_3) \tag{5}$$

The elements in (4) can be found in the Table 1.
Denotation $\mathbf{q} = [x_r, y_r, \theta_1, \theta_2, \theta_3]^T$ is internal variable vector of the system. The Lagrange equation of the system is:

$$\frac{d}{dt}\left(\frac{\partial K}{\partial \dot{\mathbf{q}}}\right) - \frac{\partial K}{\partial \mathbf{q}} + \frac{\partial P}{\partial \mathbf{q}} = \tau - \mathbf{N}^T(\mathbf{q})\lambda \tag{6}$$

where $\mathbf{N}$ is the Jacobian matrix of constrain force. λ is Lagrange multipliers, τ is the system torque vector. Replace the potential and kinetic energy into Lagrange equation (6) then reduce the Lagrange multipliers λ results in:

$$\mathbf{M}(\mathbf{q})\ddot{\mathbf{v}}_s + \mathbf{C}(\mathbf{q}, \mathbf{v}_s)\dot{\mathbf{v}}_s + \mathbf{g}(\mathbf{q}) = \tau. \tag{7}$$

where $\dot{\mathbf{v}}_s = [\dot{\phi}_R, \dot{\phi}_L, \dot{\theta}_2, \dot{\theta}_3]$ is the measured vector of angular velocity of two wheels and pan/tilt joint. The vector $\mathbf{g}(\mathbf{q}) = [0, 0, 0, m_3gl_3\sin(\theta_3)]^T$; $\mathbf{M(q)}$ is the (n x n) symmetric positive definite inertia matrix. $\mathbf{q}$ is the (n x 1) vector of integrated robot wheels and joints. $\mathbf{C}(\mathbf{q}, \mathbf{v}_s)$ is the vector of centripetal and Coriolis effects.

Assumption 2: The desired variable of the system $\mathbf{q}_d$, $\dot{\mathbf{q}}_d$, $\ddot{\mathbf{q}}_d$ and desired moment τ_d, $\dot{\tau}_d, \ddot{\tau}_d$ are exits and bounded.

3 Design Controller for the System

3.1 Design Kinematic Controller

Assumption 3: We hypothesize that the intrinsic parameters of two cameras are the same, placing at the same height and the camera's optical axis are parallel.

Notations: Left and right camera coordinates are O_L; O_R with the origin located at the focal point of the camera, camera coordinate is O_C with the origin located at the midpoint of origin of two cameras. The photo frame is specified at the front and perpendicular to the Z-axis at the center, the axis u, v parallel to the axis X, Y of camera, respectively. The feature point coordinates of the target obtained from left and right camera's image are $(U_L, V_L); (U_R, V_R)$. Following the Assumption 1, we have $V_L = V_R = V$). From Fig. 2a, the coordinates of feature point on the left and right image frame are transformed to *(X, Z)* and *(Y, Z)* plane then the coordinates of the target point *Q (X, Y, Z)* are calculated in O_C coordinates [3]:

$$^{C}\mathbf{x} = \begin{bmatrix} X \\ Y \\ Z \end{bmatrix} = \frac{1}{U_R - U_L} \begin{bmatrix} 0.5B\,(U_L + U_R) \\ BV_L \\ Bf \end{bmatrix}; \tag{8}$$

B is the distance between the optical axes of two cameras, f is the cameras focal length. $\mathbf{m} = [\mathrm{U}_L \mathrm{V}_L \mathrm{U}_R]^T, \mathbf{m}_d$ are the measured and desired image feature vector. We have the image feature error:

$$\varepsilon_i = \mathbf{m} - \mathbf{m}_d \tag{9}$$

The control law $\left[\dot{\theta}_{2d},\ \dot{\theta}_{3d}\right]^T$ is looking so $\varepsilon_i \rightarrow \mathbf{0}$. The velocity relationship between the movements of the target and it's feature image is:

$$\dot{\mathbf{m}} = \mathbf{J}_{imag}(\mathbf{m})\dot{\mathbf{x}} + \frac{\partial \mathbf{m}}{\partial t} \tag{10}$$

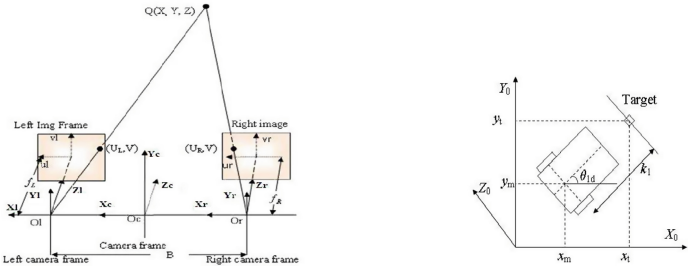
- $\dot{\mathbf{m}}$ is the velocity of the image feature vector; $\mathbf{x} = [x,\ y,\ z,\ \phi_X,\ \phi_Y,\ \phi_Z]^T; \dot{\mathbf{x}} = [T_X,\ T_Y,\ T_Z,\ \omega_X,\ \omega_Y,\ \omega_Z]^T$ are extrinsic variable vector and velocity of grip head of pan/tilt robot.
- $\frac{\partial \mathbf{m}}{\partial t}$ is the velocity of the image feature caused by the movement of target. It unknown but we can approximate by discrete sampling method.
- $\mathbf{J}_{imag}(\mathbf{m})$ is the image Jacobi matrix:

$$\mathbf{J}_{imag} = \begin{bmatrix} \frac{U_R - U_L}{B} & 0 & \frac{U_L(U_L - U_R)}{fB} & \frac{U_L V_L}{f} & -\frac{2f^2 + U_L^2 + U_L U_R}{2f} & V_L \\ 0 & \frac{U_R - U_L}{B} & \frac{V_L(U_L - U_R)}{fB} & \frac{f^2 + V_L^2}{f} & -\frac{V_L(U_L + U_R)}{2f} & -\frac{U_L + U_R}{2f} \\ \frac{U_R - U_L}{B} & 0 & \frac{U_R(U_L - U_R)}{fB} & \frac{U_R V_L}{f} & -\frac{2f^2 + U_R^2 + U_L U_R}{2f} & V_L \end{bmatrix} \tag{11}$$

Determine the desired velocity of Pan and Tilt joints
Notation the Jacobian of the mobile robot is $\mathbf{J}_r$ yields (12) and (13):

$$\dot{\mathbf{x}} = \mathbf{J}_r(\mathbf{q})\dot{\mathbf{q}} \tag{12}$$

$$\begin{bmatrix} T_X \\ T_Y \\ T_Z \\ \omega_X \\ \omega_Y \\ \omega_Z \end{bmatrix} = \begin{bmatrix} -s_{12} & -c_{12} & 0 & 0 & 0 \\ -c_{12}s_3 & -s_{12}s_3 & 0 & 0 & 0 \\ c_{12}c_3 & s_{12}c_3 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -1 \\ 0 & 0 & c_3 & c_3 & 0 \\ 0 & 0 & s_3 & s_3 & 0 \end{bmatrix} \begin{bmatrix} \dot{x}_m \\ \dot{y}_m \\ \dot{\theta}_1 \\ \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} \tag{13}$$



(a) (b)

Fig. 2. a-Stereo camera model; b-Desired position and orientation

Substituting the Eq. (7) into (5) yields:

$$\dot{\mathbf{m}} = \mathbf{J}_{imag}(\mathbf{m})\dot{\mathbf{x}} + \frac{\partial \mathbf{m}}{\partial t} = \mathbf{J}\dot{\mathbf{q}} + \frac{\partial \mathbf{m}}{\partial t} \tag{14}$$

where $\mathbf{J}$ is the general matrix:

$$\mathbf{J} = \mathbf{J}_{img}\mathbf{J}_r \tag{15}$$

The Eq. (9) shows the relationship between the velocity of image feature point and internal variable of mobile robot – pan/tilt – stereo camera system. Rewriting Eq. (9) in other form, results in:

$$\dot{\mathbf{m}} = \mathbf{J}_r \begin{bmatrix} \dot{x}_r \\ \dot{y}_r \end{bmatrix} + \mathbf{J}_p \begin{bmatrix} \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} + \mathbf{J}_1\dot{\theta}_1 + \frac{\partial \mathbf{m}}{\partial t} \tag{16}$$

$\partial \mathbf{m}/\partial t$ can be approximated according to the equation:

$$\frac{\partial \mathbf{m}}{\partial t} \approx (\dot{\mathbf{m}})^{last} - \mathbf{J}_r \begin{bmatrix} \dot{x}_r \\ \dot{y}_r \end{bmatrix} - \mathbf{J}_p \begin{bmatrix} \dot{\theta}_2^{last} \\ \dot{\theta}_3^{last} \end{bmatrix} - \mathbf{J}_1\dot{\theta}_1; \tag{17}$$

$(\dot{\mathbf{m}})^{last}; \dot{\theta}_2^{last}; \dot{\theta}_3^{last}$ is the measured value and sampling with sampling time is 0.05s. Substituting the Eq. (12) into (11) results in:

$$\dot{\mathbf{m}} = (\dot{\mathbf{m}})^{last} + \mathbf{J}_p \left(\begin{bmatrix} \dot{\theta}_2 \\ \dot{\theta}_3 \end{bmatrix} - \begin{bmatrix} \dot{\theta}_2^{last} \\ \dot{\theta}_3^{last} \end{bmatrix} \right). \tag{18}$$

The control law applied to Eq. (11) in order to $\varepsilon_i \to \mathbf{0}$ is chosen as follows:

$$\begin{bmatrix} \dot{\theta}_{2d} \\ \dot{\theta}_{3d} \end{bmatrix} = \begin{bmatrix} \dot{\theta}_2^{last} \\ \dot{\theta}_3^{last} \end{bmatrix} + \mathbf{J}_p^{+} \left(\dot{\mathbf{m}}_d - \mathbf{K}\varepsilon_i - (\dot{\mathbf{m}})^{last} \right), \tag{19}$$

$\mathbf{J}_p^{+}$ is the pseudo-inverse matrix of $\mathbf{J}_p$. $\mathbf{K}$ is diagonal, symmetric positive definite matrix. Taking the derivative of (4) and substituting $\begin{bmatrix} \dot{\theta}_2 \ \dot{\theta}_3 \end{bmatrix}^T$ into the Eq. (13) by $\begin{bmatrix} \dot{\theta}_{2d} \ \dot{\theta}_{3d} \end{bmatrix}^T$ in (14) yields:

$$\dot{\varepsilon}_i = -\mathbf{K}\varepsilon_i \tag{20}$$

Determination the speed of the mobile robot wheels

The coordinates of the target see in O_0 is ${}^0\mathbf{x}({}^0x_t, {}^0y_t, {}^0z_t)^T$ following:

$${}^0\mathbf{x} = {}^0R_C.{}^C\mathbf{x} \tag{21}$$

The desired coordinates and direction of the mobile robot is calculated by following the equation (Fig. 2b):

$$\begin{aligned} x_{rd} &= {}^0x_t - k_1 \cos\theta_b; \\ y_{rd} &= {}^0y_t - k_1 \sin\theta_b; \\ \theta_{1d} &= \theta_b = \text{atan2}({}^0\dot{y}_t, {}^0\dot{x}_t); \end{aligned} \tag{22}$$

where *k1* is the distance from the midpoint of the 2 robot's wheels to the center of the target.

The error between the current and the desire position $(x_r, y_r, \theta_1), (x_{rd}, y_{rd}, \theta_{1d})$ of mobile robot is defined as:

$$\varepsilon_r = \begin{bmatrix} \varepsilon_{r1} \\ \varepsilon_{r2} \\ \varepsilon_{r3} \end{bmatrix} = \begin{bmatrix} \cos\theta_1 & \sin\theta_1 & 0 \\ -\sin\theta_1 & \cos\theta_1 & 0 \\ 0 & 0 & 1 \end{bmatrix} \begin{bmatrix} x_{rd} - x_r \\ y_{rd} - y_r \\ \theta_{1d} - \theta_1 \end{bmatrix}. \tag{23}$$

Taking the derivative of ε_r with $\omega_{rd} = \dot{\theta}_{1d}, v_{rd} = \sqrt{\dot{x}_{rd} + \dot{y}_{rd}}$, yields:

$$\dot{\varepsilon}_r = \begin{bmatrix} \dot{\varepsilon}_{r1} \\ \dot{\varepsilon}_{r2} \\ \dot{\varepsilon}_{r3} \end{bmatrix} = \begin{bmatrix} \omega_r \varepsilon_{r2} - v_r - v_{rd} \cos\varepsilon_{r3} \\ -\omega_r \varepsilon_{r1} + v_{rd} \sin\varepsilon_{r3} \\ \omega_{rd} - \omega_r \end{bmatrix} \tag{24}$$

Backstepping method is used to build a kinematic controller to track a moving target with desire velocity. Here, the straight and angular velocity of mobile robot is chosen following [3] as:

$$\begin{bmatrix} v_m \\ \omega_m \end{bmatrix} = \begin{bmatrix} v_{rd} \cos\varepsilon_{r3} + k_2 \varepsilon_{r1} \\ \omega_{rd} + k_3 \varepsilon_{r3} + v_{rd} \varepsilon_{r2} \frac{\sin\varepsilon_{r3}}{\varepsilon_{r3}} \end{bmatrix} \tag{25}$$

Substituting (19) into (18) result in:

$$\dot{\varepsilon}_r = \begin{bmatrix} \dot{\varepsilon}_{r1} \\ \dot{\varepsilon}_{r2} \\ \dot{\varepsilon}_{r3} \end{bmatrix} = \begin{bmatrix} \omega_m \varepsilon_{r2} - k_1 \varepsilon_{r1} \\ -\omega_m \varepsilon_{r1} + v_d \sin \varepsilon_{r3} \\ -k_3 \varepsilon_{r3} - v_d \varepsilon_{r2} (\sin \varepsilon_{r3} / \varepsilon_{r3}) \end{bmatrix} \tag{26}$$

where k_2, k_3 are positive coefficient. If $\varepsilon_{r3} \to 0$, $(\sin \varepsilon_{r3})/\varepsilon_{r3} \to 1$, then ω_m is bounded. Following Assumption 2, we have the desired angular velocity of the two wheels of mobile robot as in Eq. (27):

$$\begin{bmatrix} \dot{\phi}_{rightd} \\ \dot{\phi}_{leftd} \end{bmatrix} = \begin{bmatrix} r/2 & r/2 \\ r/k & -r/k \end{bmatrix}^{-1} \begin{bmatrix} v_m \\ \omega_m \end{bmatrix} \tag{27}$$

From kinematic controller (15) and (22), we can be proved the asymptotic stability of the system with candidate Lyapunov function as follows:

$$V_k = \frac{1}{2} \left(\varepsilon_i^T \varepsilon_i + \varepsilon_r^T \varepsilon_r \right) \tag{28}$$

3.2 Design Adaptive Controller for Dynamic System

Now we are going to design an adaptive controller for dynamic system, which showed in Eq. (7). Notation the desired vector of angular velocity of two wheels and pan/tilt joint is $\dot{\mathbf{v}}_d = [\dot{\phi}_{Rd}, \dot{\phi}_{Ld}, \dot{\theta}_{2d}, \dot{\theta}_{3d}]$ The tracking error between the measured velocity and desired velocity of the state variables denote as:

$$\mathbf{e} = \mathbf{v}_d - \mathbf{v}_s. \tag{29}$$

Using Adaptive controller with Computed-Torque-Control method, the momentum applied for pan/tilt joints and two wheels of mobile robot are choose as:

$$\tau = \hat{\mathbf{M}}(\mathbf{q})(\ddot{\mathbf{v}}_d + \mathbf{K}_v \dot{\mathbf{e}} + \mathbf{K}_p \mathbf{e}) + \hat{\mathbf{C}}(\mathbf{q}, \mathbf{v}_s)\dot{\mathbf{v}}_d + \hat{\mathbf{g}}(\mathbf{q}) \tag{30}$$

$\hat{\mathbf{M}}(\mathbf{q})$, $\hat{\mathbf{C}}(\mathbf{q}, \mathbf{v}_s)$, $\hat{\mathbf{g}}(\mathbf{q})$ are the estimated values of each item in the dynamic equation (7). We can obtain:

$$\hat{\mathbf{M}}(\mathbf{q})\ddot{\mathbf{v}}_s + \hat{\mathbf{C}}(\mathbf{q}, \mathbf{v}_s)\dot{\mathbf{v}}_s + \hat{\mathbf{g}}(\mathbf{q}) = \mathbf{D}\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{v}}_s\right) \hat{\mathbf{a}}. \tag{31}$$

where $\mathbf{D}\left(\mathbf{q}, \dot{\mathbf{q}}, \dot{\mathbf{v}}_s\right)$ is a (n × r) regressor matrix. $\hat{\mathbf{a}} = [I_{1z}, I_{2z}, I_{3x}, I_{3y}, I_{3z}]^T$ is the vector that summarizes all the estimated parameters. Substituting the control τ (30) into the dynamic equation (7) result the closed-loop error model:

$$\hat{\mathbf{M}}(\mathbf{q})(\ddot{\mathbf{e}} + \mathbf{K}_v \dot{\mathbf{e}} + \mathbf{K}_p \mathbf{e}) = \mathbf{D}\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{v}}_s\right) \tilde{\mathbf{a}}, \tag{32}$$

where $\tilde{\mathbf{a}} = \hat{\mathbf{a}} - \mathbf{a}$. Notation $\mathbf{s} = \begin{bmatrix} \mathbf{e}^T & \dot{\mathbf{e}}^T \end{bmatrix}^T$ and rewriting the Eq. (36) yields:

$$\dot{\mathbf{s}} = \mathbf{A}\mathbf{s} + \mathbf{B}\hat{\mathbf{M}}^{-1}(\mathbf{q})\mathbf{D}\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{v}}_s\right) \tilde{\mathbf{a}}, \tag{33}$$

where $\mathbf{A} = \begin{bmatrix} \mathbf{0}_{n\,x\,n} & \mathbf{I}_{n\,x\,n} \\ -\mathbf{K}_p & -\mathbf{K}_v \end{bmatrix}; \mathbf{B}(\mathbf{q}) = \begin{bmatrix} \mathbf{0}_{n\,x\,n} \\ \mathbf{I}_{n\,x\,n} \end{bmatrix}$. The adaptive control law is design as:

$$\dot{\hat{\mathbf{a}}} = -\boldsymbol{\Gamma}^{-1}\mathbf{D}^T\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{v}}_s\right)\hat{\mathbf{M}}^{-1}(\mathbf{q})\mathbf{B}^T\mathbf{P}\mathbf{s} \tag{34}$$

where Γ is an *(r x r)* diagonal positive - definite constant matrix, P is a*(2n x 2n)* symmetric positive - definite constant matrix satisfying

$$\mathbf{A}^T\mathbf{P} + \mathbf{P}\mathbf{A} = -\mathbf{Q} \tag{35}$$

Q is a symmetric positive - definite constant matrix (Figs. 3 and 5)

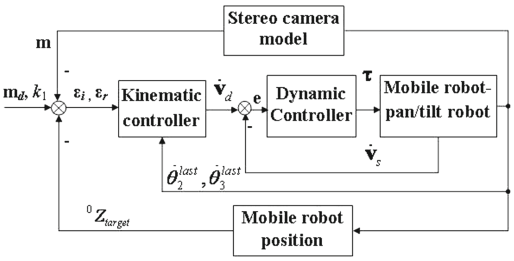


Fig. 3. Structure of the system

3.3 Stability of the System

We choose the candidate Lyapunov function as follows:

$$V(\mathbf{s}, t) = \frac{1}{2}\mathbf{s}^T(t)\mathbf{P}\mathbf{s}(t) + \frac{1}{2}\tilde{\mathbf{x}}^T\boldsymbol{\Gamma}^{-1}\tilde{\mathbf{a}}; \tag{36}$$

Because **P**, **Γ** are a symmetric positive definite matrix, then $V > 0$ when $\mathbf{s} \neq \mathbf{0}$; $V = 0$ if only if $\boldsymbol{s} = 0$. Taking the derivative both sides of V along time, yields:

$$\dot{V} = \frac{1}{2}\dot{\mathbf{s}}^T\mathbf{P}\mathbf{s} + \frac{1}{2}\mathbf{s}^T\mathbf{P}\dot{\mathbf{s}} + \tilde{\mathbf{x}}^T\boldsymbol{\Gamma}^{-1}\dot{\tilde{\mathbf{a}}}. \tag{37}$$

Because **Γ** is a diagonal positive - definite constant matrix:

$$\boldsymbol{\Gamma} = \boldsymbol{\Gamma}^T;\ \left[\tilde{\mathbf{x}}^T\boldsymbol{\Gamma}^{-1}\dot{\tilde{\mathbf{a}}}\right]^T = \tilde{\mathbf{a}}^T\boldsymbol{\Gamma}^{-1}\dot{\tilde{\mathbf{a}}} \tag{38}$$

Substituting $\dot{\mathbf{s}}$ from Eq. (33) into (37) and using the Eq. (35) yields:

$$\dot{V} = -\frac{1}{2}\mathbf{s}^T\mathbf{Q}\mathbf{s} + \tilde{\mathbf{a}}^T(\hat{\mathbf{M}}^{-1}(\mathbf{q})\mathbf{D}\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{v}}_s\right)\mathbf{B}^T\mathbf{s} + \boldsymbol{\Gamma}^{-1}\dot{\tilde{\mathbf{a}}}). \tag{39}$$

By choice the adaptation update rule is (40). The Eq. (39) become (41):

$$\dot{\tilde{\mathbf{a}}} = -\boldsymbol{\Gamma}^{-1}\hat{\mathbf{M}}^{-1}(\mathbf{q})\mathbf{D}\left(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{v}}_s\right)\mathbf{B}^T\mathbf{s} \tag{40}$$

$$\dot{V} = -\frac{1}{2}\mathbf{s}^T\mathbf{Q}\mathbf{s}. \tag{41}$$

Because $\mathbf{Q}$ is a symmetric positive definite matrix, from Eq. (41) we have:

$$\dot{V} \leq -\frac{1}{2}\mathrm{w}_1\|\mathbf{s}\|^2 \leq 0. \tag{42}$$

where w_1 is the smallest value of $\mathbf{Q}$. The adaptive update rule is.

$$\dot{\hat{\mathbf{a}}} = \mathbf{\Gamma}\hat{\mathbf{M}}^{-1}(\mathbf{q})\mathbf{D}^T(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{v}}_s)\mathbf{B}^T\mathbf{s} \tag{43}$$

We see that V is positive definite function. $\dot{V} \leq 0$ is negative-semidefinite function. But, to determine explicitly the type of stability of the system we going to calculate the second derivative of V:

$$\ddot{V} = -\mathbf{s}^T\mathbf{Q}\dot{\mathbf{s}}\,. = -\mathbf{s}^T\mathbf{Q}(\mathbf{A}\mathbf{s} + \mathbf{B}\hat{\mathbf{M}}^{-1}(\mathbf{q})\mathbf{D}(\mathbf{q}, \dot{\mathbf{q}}, \ddot{\mathbf{v}}_s)\tilde{\mathbf{a}}) \tag{44}$$

Following the Assumption 2, $\mathbf{q}_d$, $\dot{\mathbf{q}}_d$, $\ddot{\mathbf{q}}_d$ are bounded. Therefore, we have $\mathbf{s}$, $\dot{\mathbf{s}}$, $\tilde{\mathbf{x}}$ bounded. In other hand, $\mathbf{Q}$ is a symmetric positive definite matrix. So According to Barbalat's lemma, $\dot{V}$ is uniformly continuous and $\ddot{V}$ is bounded, $\mathbf{e} \to \mathbf{0}$, $\dot{\mathbf{e}} \to \mathbf{0}$ when $t \to \infty$ the system is asymptotically stable. The tracking error will asymptotic to $\mathbf{0}$

4 Simulation

In simulations, we assumpt that mobile robot following to track the target and far away from the target a distance $k1 = 0.4\,\mathrm{m}$ in two axes x, y.

$\mathbf{\Gamma} = diag\,[5; 5; 5; 5; 5]\,; \mathbf{K}p = diag\,[100; 100]\,; \mathbf{K}v = diag\,[10; 10]\,; \mathbf{P} = diag\,[2; ...; 2]$

Table 1. The parameters of the system

Parameter name	Notation	Value
The distance $l1$	$l1$	0.1 m
Length of pan joint; tilt joint	$l2$; $l3$	0.22 m; 0.1 m
Mass of mobile robot; wheel	$m_1; m_w$	23 kg; 1 kg
Mass of pan joint; tilt joint	$m_2; m_3$	2.2 kg; 0.8 kg
Inertial moment of the wheel	I_w	$0.005\,kgm^2$
Inertial moment of mobile robot in z axis	I_1z	$1.565\,kgm^2$
Inertial moment of pan joint in z axis of O_2	I_2z	$0.0124\,kgm^2$
Inertial moment of tilt joint in x; y; z axes O_3	$I_3x; y; z$	$0.0124; 0.0146; 0862\,kgm^2$
Distance of 2 cameras; wheel radius	K; r	0.2 m; 0.2 m
Focal length of the camera (Eye-RIS 2.1)	f	12.5 mm

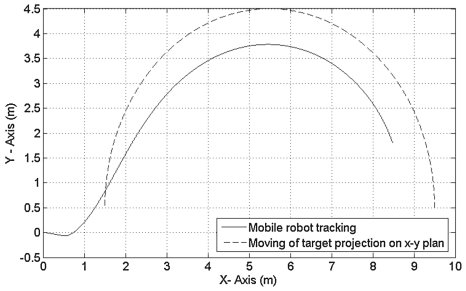


Fig. 4. Tracking of mobile robot

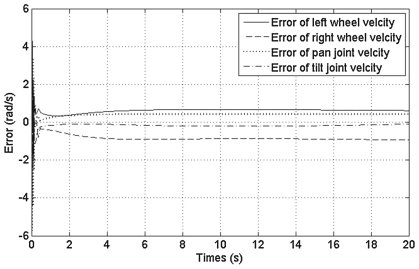


Fig. 5. Error between desired and measured angular velocity

Simulation: Moving target in an arc: Moving the target follows the circular arc with center coordinates at the origin O(5.5 m, 0.5 m), radius r = 4. Target moves from point A(1.5, 0.5, 0)m to B(9.0, 0.5, 1)m following arc in O_0 coordinates with time T = 20 s. Simulation results are given in Fig. 4.

5 Conclusions

In this paper, we constructed the kinematic and dynamic models for the integrated system. We also design an adaptive control for the system to estimate the uncertain parameters. However, the simulation resulting are not so well and still have a pseudo inverse matrix then we will research about inverse property of it in other paper.

References

1. Tinh, N.V., Cat, P.T., Tuan, P.M., Quyen Bui, T.T.: Visual control of integrated mobile robot-pan tilt-camera system for tracking a moving target. In: IEEE Robio 2014, pp. 1566–1571 (2014)
2. Thuan, T.H., Tinh, N.V., Tuan, P.M., Cat, P.T.: Trajectory tracking control of a mobile robot by computed torque method with on-line learning neural network. In: IEEE 2013, pp. 1184–1189 (2013)

3. Chung, L.V., Cat, P.T.: A new control method for stereo visual servoing system with pan-tilt platform. Comput. Sci. Cybern. **31**(2), 107–122 (2015)
4. Lupián, L.F., Rabadán-Martin, J.R.: LQR control methods for trajectory execution in omnidirectional mobile robots. In: Topalov, A.V. (ed.) Recent Advances in Mobile Robotics, pp. 385–401. InTech, Rijeka (2008)
5. Gutiérrez-Arias, E.M., Flores-Mena, J.E., Morin-Castillo, M.M., Suárez-Ramírez, H.: Design of an optimal control for an autonomous mobile robot. Rev. mexicada F'is. **57**, 75–83 (2011)
6. Baroudi, M., Saad, M., Ghie, W.: State-feedback and linear quadratic regulator applied to a single-link flexible manipulator. In: IEEE Robio 2009, pp. 1381–1386 (2009)
7. Fang, J.: The LQR controller design of two-wheeled self-balancing robot based on the particle swarm optimization algorithm. Math. Probl. Eng. **2014**, 6 (2014). Article ID: 729095
8. Berg, J.V.D., David, W., Stephen, J., Marc, N., Dinesh, M.: LQG-obstacles feedback control with collision avoidance for mobile robots with motion and sensing uncertainty. In: ICRA 2012, pp. 346–353 (2012)
9. Kabanov, A.A.: Optimal control of mobile robot's trajectory movement. WSEAS Trans. Syst. Control **9**, 398–404 (2014)
10. Widodo B., Ari S., Djoko P., Achmad J.: A navigation system for service robot using stereo vision and kalman filtering. In: IEEE 2011, pp. 1771–1776 (2011)
11. Lewis, F.L., Syrmos, V.L.: Optimal Control. Wiley, New York (1995)

Estimation Localization in Wireless Sensor Network Based on Multi-objective Grey Wolf Optimizer

Trong-The Nguyen[1,2], Ho Thi Huong Thom[1,2], and Thi-Kien Dao[1,2(✉)]

[1] Department of Information Technology,
Hai Phong Private Universty, Haiphong, Vietnam
vnthe@hpu.edu.vn, jvnkien@gmail.com
[2] Faculty of Information Technology,
Vietnam Maritime University, Haiphong, Vietnam
thomhth@vimaru.edu.vn

Abstract. Determining the position of nodes of a network plays an important role in many wireless sensor networks (WSN) applications e.g. in tracking, detecting, monitoring, etc. In this paper, the multi-objective grey wolf optimizer (MGWO) for the estimating approaches of the located nodes in a network is proposed to solve the multi-objective optimization localization issues in WSNs. There two objective functions related to the estimation localization are the distance of nodes and the geometric topology that consider to formula multiobjective optimization localization. The simulation results show considerable improvements in terms of localization accuracy and convergence rate in comparison with those obtained from the other methods.

Keywords: Multi-objective grey wolf optimizer · Swarm intelligent · Node localization · Wireless sensor networks

1 Introduction

A wireless sensor network (WSN) could consist of hundreds or even thousands of low-cost nodes communicating among themselves [1]. It has become an important technology especially for military applications, disaster management, wildlife and environmental monitoring [2]. In applications of WSN, such as environment monitoring, precision agriculture, vehicle tracking, and logistics, knowledge about the location of sensor nodes plays a key role [3]. The correlation of sensor measurements with physical locations is required in these applications, even if the accessible knowledge about positions of nodes is only approximate. Moreover, information about current locations are used in geographical-based routing, data aggregation and various network services. Hence, self-organization and localization capabilities are one of the most important requirements in sensor networks. Theoretically, location awareness can be enabled in principle by the use of a Global Positioning System (GPS). However, this solution is not always viable in

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_25

practice, because a sensor network consists of thousands of nodes and GPS will be very costly. In addition, GPS is not well suited to indoor and underground deployments, and the presence of obstacles like dense foliage or high buildings may impair the outdoor communication with satellites. Several alternative techniques have been developed to deal with these limitations, as reviewed in [3,4], among which fine-grained localization techniques may represent the most suitable ones. Instead of requiring all nodes installed GPS, in these schemes, only a few nodes of the network are called reference or anchor nodes which are endowed with their exact positions through GPS or manual placement. While other nodes in a network are able to derive their positions by estimating their distances to nearby nodes with using the measurement techniques included received signal strength (RSS) measurements, and time of arrival. Traditionally, most studies focused on using single-objective optimization problem to solve localization problems. The localization problem was modeled as a single-objective function for optimization with the space distance constraint. These studies have also achieved significant results in both accuracy and computational time. However, the single objective function did not count really all impacting or affecting from the other objective factors. But, if there is the combination of the objective functions into a multiobjective function that will obtain the meaningful result. For example, in some applications, the obtained results of estimated nodes localizations could meet the space distance constraint, but they could not meet the geometric topology constraint due to ranging errors. Recently, some works have proved the effectiveness of multiobjective optimization algorithms to solve conflict multiple objectives [6]. It is more reasonable to model the node localization as a multiobjective optimization problem, which can be described as solving a Pareto solution, rather than simply being described as a single-objective optimization problem. Based on this viewpoint, in this paper, a multiobjective model is adopted to solve the node localization problem with objective functions included the distances constraint and the topological constraint. Pareto optimal solutions for obtaining optimal solution is achieved by applying the multiobjective Grey wolf optimizer (MGWO).

2 Localization Model

WSN assumes with n nodes that are deployed in two-dimensional space of Z^2 including m anchor nodes and $n\breve{}m$ unknown nodes in which $m < n$. The objective localization in a WSN is to estimate the coordinates of $n\breve{}m$ unknown nodes using the a priori information about the location of m anchor nodes. The coordinates of unknown nodes need to meet both the space distance constraint and the geometric topology constraint. An example of localization model has two objective functions that included the space distances and the geometric topology. The reason for meeting the constraint of space distance is to make the estimated coordinates close to the real values, and the reason of meeting the geometric topology constraint is to make the network topology unique [3]. In the space distance constraint, objective function for the WSN localization includes two-phase process. In the first phase, it was known as ranging process which nodes

estimate their distances from anchor nodes using the signal propagation time or the received signal strength indicator (RSSI). In the second phase, position estimation of the nodes is carried out using the ranging information [7]. The localization error is minimized by using the optimization algorithm. Supposing the two nodes i and j being in the communication radius of each other and effect of measurement noise is simulated as a Gaussian additive white noise. In the first phase, each anchor nodes in the deployment estimates its distance from each of its neighboring target nodes. The internode ranging distance can be obtained by RSSI ranging technology and denoted as following.

$$d_{ij} = r_{ij} + n_{ij} \tag{1}$$

where r_{ij} is the actual distance between two nodes and n_{ij} is a ranging error. r_{ij} is able to be calculated as given.

$$r_{ij} = \sqrt{((x_i - x_j)^2 + (y_i - y_j)^2)} \tag{2}$$

where (x_i, y_i) and (x_j, y_j) are coordination of node i and j location. The neighbor effect factor N_i is set to $j \in 1, .., n, j \neq i$ if $r_{ij} \leq R$, and its complement (N_i) is set to $j \in 1, .., n, j \neq i$ if $r_{ij} \leq R$. Where R is the communication of node i. The measurement noise n_{ij} is the ranging error of RSSI which follows a zero mean Gaussian distribution with variance σ^2, and it has a random value uniformly distributed in the range $[d_i - d_i \frac{P_n}{100}, d_i + d_i \frac{P_n}{100}]$. In second phase, the objective function for the space distance constraint can be framed as.

$$f_1 = \sum_{i=m+1}^{n} (\sum_{j \in N_i} (\widehat{d_{ij}} - d_{ij})^2) \tag{3}$$

where $(\hat{d_{ij}})$ is the estimated distance between nodes i and j, defined by.

$$\widehat{d_{ij}} = \begin{cases} \sqrt{(\widehat{x_i} - x_j)^2 + (\widehat{y_i} - y_j)^2} & if\ j\ is\ an\ anchor, \\ \sqrt{(\widehat{x_i} - \widehat{x_j})^2 + (\widehat{y_i} - \widehat{y_j})^2} & otherwise \end{cases} \tag{4}$$

The second objective function of the geometric topology constraint is defined by.

$$f_2 = \sum_{i=m+1}^{n} (\sum_{j \in N_i} \theta_{ij} + \sum_{j \in \overline{N_i}} (1 - \theta_{ij})) \tag{5}$$

The geometric topology constraint represents the connectivity constraint which dissatisfies the current estimated positions of non-anchor nodes [7]. And θ_{ij} is denoted by

$$\theta_{ij} = \begin{cases} 1 & if\ \hat{d}_{ij} > R \\ 0 & otherwise \end{cases} \tag{6}$$

The space distance constraint and the geometric topology constraint imply the accuracy of the coordinates of the nodes. The high accuracy of estimated coordinates of the unknown nodes consequently lead to the small values of the two

objective functions. Therefore, estimating coordinates of the unknown nodes can be modeled to search the optimum solution for the multiobjective optimization issues, which can be obtained by decreasing both values of the objective functions f_1 and f_2.

3 Multiobjective Grey Wolf Optimization for Localization

Basic version of the grey wolf optimizer (GWO) is only for single objective optimization. In order to solve multiobjective functions of the localization in WSN, GWO is extended to multiobjective grey wolf optimizer (MGWO). The basic version of GWO and Pareto optimal front are first briefly reviewed, and the localization problem in WSN will be dealt with it based on MGWO.

3.1 The Basic Grey Wolf Optimizer

GWO is inspired from modeling the leadership hierarchy and hunting mechanism of Grey wolf when searching and attacking for the prey [8]. There are four guided types of grey wolves in the leadership hierarchy, such as alpha (α), beta (β), delta (δ), and omega (ω). Type of α is considered the fittest solution, and then β, and δ are considered the second and the third best solutions respectively. Omega (ω) could be assumed the rest of the candidate solutions. GWO algorithm consists of the constructed mathematical models as follows: The dominance degree in the social leadership hierarchy is formulated in Equations of model as follows:

$$\boldsymbol{D} = \left| \overrightarrow{C} \ \overrightarrow{X_p}(t) - \ \boldsymbol{X}(t) \right| \tag{7}$$

$$\boldsymbol{X}(t+1) = \overrightarrow{X_p}(t) - \ \boldsymbol{A}\boldsymbol{D} \tag{8}$$

where $\overrightarrow{D}$ is dominance degree, t indicates the current iteration, $\overrightarrow{A}$ and $\overrightarrow{C}$ are coefficient vectors, $\overrightarrow{(X_p)}(t)$ is the position vector of the prey, and $\overrightarrow{X(t)}$ indicates the position vector of a grey wolf. Equations (1) and (2) are two-dimensional position vector and some of the possible neighbors. The vectors $\overrightarrow{A}$ and $\overrightarrow{C}$ are calculated as follows:

$$\boldsymbol{A} = 2\boldsymbol{a}\overrightarrow{r_1} - \ \boldsymbol{a} \tag{9}$$

$$\boldsymbol{C} = 2\overrightarrow{r_2} \tag{10}$$

where components $\overrightarrow{a}$ are linearly decreased from 2 to 0 over the course of iterations and r1, r2 are random vectors in [0, 1]. A grey wolf in the position of (X, Y) can update its position according to the position of the prey (X^*, Y^*). Different places around the best agent can be reached with respect to the current position by adjusting the value of $\overrightarrow{A}$ and $\overrightarrow{C}$ vectors. The hunting behavior of grey wolves can be simulated when the alpha (best candidate solution), beta, and delta are supposed to have better knowledge about the potential location of prey. Therefore, the first three best solutions are obtained so far and oblige the

other search agents (including the omegas) to update their positions according to the best search agents. This simulating model is formulated as follows:

$$\overrightarrow{D_\alpha} = \left|\overrightarrow{C_1}\ \overrightarrow{X_\alpha}\ - \boldsymbol{X}\right|, \overrightarrow{D_\beta} = \left|\overrightarrow{C_2}\ \overrightarrow{X_\beta}\ - \boldsymbol{X}\right|, \overrightarrow{D_\delta} = \left|\overrightarrow{C_3}\ \overrightarrow{X_\delta}\ - \boldsymbol{X}\right| \tag{11}$$

$$\overrightarrow{X_1} = \overrightarrow{X_\alpha} - \overrightarrow{A_1}\left(\overrightarrow{D_\alpha}\right), \overrightarrow{X_2} = \overrightarrow{X_\beta} - \overrightarrow{A_2}\left(\overrightarrow{D_\beta}\right), \overrightarrow{X_3} = \overrightarrow{X_\delta} - \overrightarrow{A_3}\left(\overrightarrow{D_\delta}\right) \tag{12}$$

$$\boldsymbol{X}\left(t+1\right) = \frac{\overrightarrow{X_1} + \overrightarrow{X_2} + \overrightarrow{X_3}}{3} \tag{13}$$

The position of the prey is estimated by alpha, beta, and delta and other wolves update their positions randomly around the prey during the hunt. The grey wolves finish the hunt by attacking the prey when it stops moving.

3.2 Pareto Optimal Front

The domination of a solution vector $x = (x_1, x_2, .., x_n)^T$ on a vector $y = (y_1, y_2, .., y_n)^T$ for a minimization problem if and only if $x_i \leq y_i$ for $\forall_i \in 1, ..., n$ and $\exists_i \in 1, ..., n : x_i < y_i$. It means that is no component of x is larger than the corresponding component of y, and at least one component is smaller. Similarly, the dominance relationship could be defined by.

$$x \preccurlyeq y \ \Leftrightarrow x \ \prec y \ \vee x = y \tag{14}$$

For maximization problems, the dominance can be defined by replacing symbol of $\prec$ with the symbol of $\succ$. Therefore, a point x_* is called a non-dominated solution if no solution can be found that dominates on it. The Pareto front PF of a multiobjective can be defined as the set of non-dominated solutions as following.

$$PF \ = \ \{s \ \in \ S | \nexists \ s' \in \ S : s' \prec \ s\} \tag{15}$$

where S is the solution set. A good approximation could be obtained from the Pareto front, if a diverse range of solutions should be generated using efficient techniques [9].

3.3 Optimal Localization Based on MGWO

There two new components are integrated in MGWO [10]. The first one is an archive, which is responsible for storing non-dominated Pareto optimal solutions obtained so far. The second component is a leader selection strategy that assists to choose alpha, beta, and delta solutions as the leaders of the hunting process from the archive. The archive is a simple storage unit that can save or retrieve non-dominated Pareto optimal solutions obtained so far. An archive controller is used to control the archive when a solution wants to enter the archive or when the archive is full. Non-dominated solutions obtained so far are compared against the archive residents during the course of iteration. The optimal solution of multiobjective optimization can be obtained from the Pareto optimal solution.

Multiobjective optimization issue for a minimization problem with d-dimensional decision vectors and h objectives is given by.

$$Minimize\ F(x) = (f_1(x),\ f_2(x),\ ..f_h(x)),\ Subject\ to\ \ x \in [x_L, x_U] \quad (16)$$

here x is a decision vector as a set of $(x_1, x_2, \ldots, x_u) \in X \in R^d$ and $F(x)$ is the objective function with the objective vector as a set of $(f_1, f_2, \ldots, f_u) \in Y \in R^h$. The decision vector x is belonging to the d-dimensional decision space X, which is corresponding to the space d dimensional of wolves in GWO. The objective function $F(x)$ belongs to the h-dimensional objective space Y, in which it is mapping functions from the decision space to the objective space. x_L, and x_U are lower and upper bound constraints of the agent range, respectively. The set of all the wolves meeting the constraints forms the decision space feasible set $\Omega = x \in R^d | x \in [x_L, x_U]$. An appropriate definition of objective functions with associated non-linear constraints. Once the tolerance or a fixed number of iterations is defined, the iterations start with the evaluation of brightness or objective values of all the wolves and compare each pair of wolves. Then, a random weight vector is generated, so that a combined best solution can be obtained. Noticed, the sum of generated weight equal to 1. The non-dominated solutions are then passed onto the next iteration. At the end of a fixed number of iterations, the non-dominated solution points can be obtained to approximate the true Pareto front. The random numbers of the weight are generated for optimization of a combined objectives functions are summed up as.

$$F(x) = w_k f_1 + (1 - w_k) f_2, \sum_{k=1}^{P} w_k = 1 \quad (17)$$

Here w_k is the weight which is generated by p_k/P, where p_k are random numbers, and P is a re-scaling operation that is generated uniformly. In the leader selection mechanism, three of the best solutions obtained so far are used as alpha, beta, and delta wolves. These leaders guide the other search agents toward promising regions of the search space with the hope to find a solution closes to the global optimum. As mentioned above there is an archive of the best non-dominated solutions obtained so far. The leader selection component chooses the least crowded segments of the search space and offers one of its non-dominated solutions as alpha, beta, or delta wolves. The weights of choosing a hypercube to select leaders from is increased when the number of obtained solutions is decreased in the hypercube. MGWO is avoided from picking similar leaders for alpha, beta, or delta. Consequently, the search is always toward the unexplored/unexposed areas of the search space since the leader selection mechanism favors the least crowded hypercube and offers leaders from different segments if there is not enough number of leaders in the least crowded segment. The employed external archive effectively saves the best non-dominated solutions obtained so far. The main steps of the algorithm process are shown in Fig. 2 as the pseudo code of multiobjective Grey Wolf Optimizer (MGWO) for localization in WSN.

1. Define objective function $F(x) = f_1(x), f_2(x)$
   where $x = (x_1, ..., x_d)^T$
2. Initialize a population of N wolves $x_i, (i = 1, 2, ..., n)$
3. **while** ($t <$ MaxGeneration)
   for $i, j = 1$ *to* N
   Evaluate their approximations PF_i and PF_j to the Pareto front
   if $i \neq j$ and when all the constraints are satisfied
4. **if** PF_j dominates PF_i,
   Optimize with GWO
   Generate new ones if the moves do not satisfy all the constraints
   end if
5. **if** no non-dominated solutions can be found
   Generate random weights w_k, $(k = 1, \ldots, P)$
   Find the best solution (among all wolves) to minimize in Eq. (17)
   end if
6. Update and pass the non-dominated solutions to next iterations
   end for
7. Sort and find the current best approximation to the Pareto front
8. Update $t = t + 1$
   end while
   Post process results and visualization;

Fig. 1. OPseudo code of Multiobjective Grey Wolf Optimizer (MGWO) for localization.

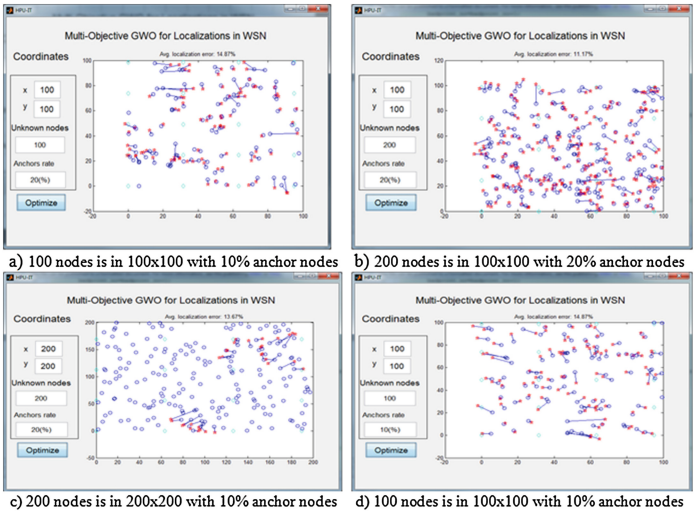


Fig. 2. Setting up a network with different scenarios included: (a) the number of nodes was 100 in an area of 100 m × 100 m with 10 % anchor node rate; (b) the number of nodes was 200 in an area of 100 m × 100 m with 20 % anchor node rate; (c) the number of nodes was 200 in an area of 200 m× 200 m with 20 % anchor node rate; and (d) the number of nodes was 1100 in an area of 100 m × 100 m with 10 % anchor node rate.

4 Simulations and Analysis

In this section, the estimation of unknown nodes for the optimal localization in a sparse network based on the multiobjective Grey Wolf Optimizer (MGWO) method is investigated. The simulations have been done applying MGWO with objective functions f_1 and f_2, and the simulation results are compared with the obtained from the original GWO and the Pareto archived evolution strategies (PAES) [12] methods. In these simulations, we focus on the average localization errors rate. To evaluate the proposed method MGWO, the different situations have been implemented such as in varying nodes densities, anchor nodes, and diversity of the Pareto solutions for optimization localization. Figure 2 shows the results of the proposed method for estimation of unknown nodes.

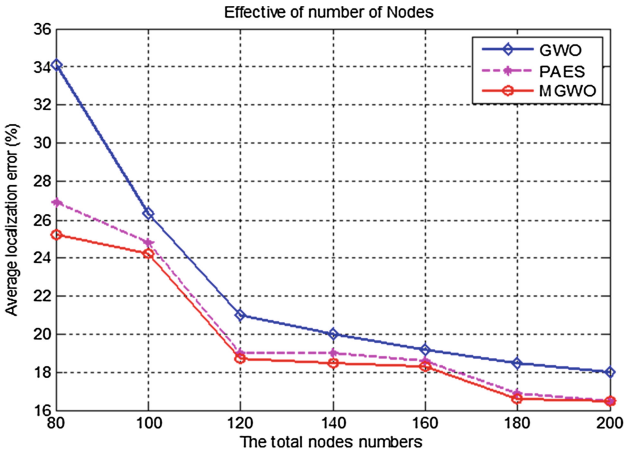


Fig. 3. Comparison of the effective density to localization errors.

Simulation results show the effectiveness of the proposed two objective functions in tackling the fine-grained localization problem in WSNs. Sensor localization for the whole sensor network was conducted in the following manner. The network consists of total n nodes are randomly deployed in specified area e.g. $100\,\mathrm{m} \times 100\,\mathrm{m}$ with m anchor nodes being randomly generated from these nodes. Assuming that the RSSI ranging error e_{ij} follows a Gaussian distribution. The different scenarios of setting the network were the variety of percentage of anchor nodes and the density of nodes in the network, that shows in Fig. 2.

Figure 3 reports the average localization errors, measured under the condition of changing the network nodes density and the total number of nodes, while holding on the anchor node proportion as 20 % and the communication radius of the radio range is fixed at 25 m. The increasing the number of anchors and the radio range of communication usually make the localization result more satisfactory, but this also implies incurring more cost.

Table 1. The effect of the proportion for localization errors with different anchor nodes rate

Methods	Anchor node proportion					
	5 %	10 %	15 %	20 %	25 %	30 %
GWO	32.20 %	26.30 %	24.00 %	21.00 %	20.20 %	19.50 %
PAES [11]	24.80 %	22.10 %	19.00 %	19.00 %	18.60 %	16.60 %
MGWO	24.20 %	21.20 %	18.70 %	18.40 %	18.20 %	16.55 %

Table 1 shows the relationship between the average localization errors and the anchor node proportion while keeping for setting the of communication radius of the fixed radio range to 25 m. All the average localization errors decrease as the anchor node proportion increases due to the increasing of anchor nodes around unknown nodes resulting in localization accuracy being improved.

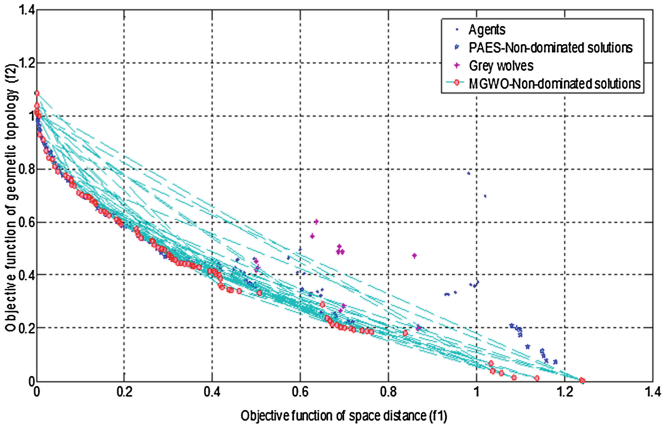


Fig. 4. The Pareto solutions of the PAES and MGWO method for localization WSN.

Compared with PAES, the proposed method MGWO reduces the average localization error 0.95 % and 0.06 %, respectively under the condition of 10% and 20 % anchor node proportion. The representative set of Pareto optimal solutions were obtained from the proposed method MGWO and PAES method that shown in Fig. 4. Visibly, the optimal solutions of the proposed method of MGWO is closer to Pareto front then the obtained of PAES method. It means that the proposed method of MGWO is more accuracy than that PAES method.

5 Conclusive

In this paper, we proposed an optimization localization in wireless sensor network (WSN) based on multiobjective Grey wolf optimizer (MGWO). The localization model has made up two objective functions included the space distance

constraint and the geometric topology constraint. The simulation results were compared with the obtained of PAES method, the localization accuracy of the proposed method MGWO is increased and the diversity rate of the proposed method is better than PAES method. Compared with traditional GWO localization algorithm, the proposed method of MGWO can improve the localization accuracy and convergence rate. In generally, the performance of the proposed method is an alternative method as a reasonable reference.

References

1. Akyildiz, I.F., Su, W., Sankarasubramaniam, Y., Cayirci, E.: A survey on sensor networks. IEEE Commun. Mag. **40**, 102–105 (2002)
2. Li, Y., Thai, M.T.: Wireless sensor networks and applications. Springer Science & Business Media, New York (2008)
3. Tepedelenlioglu, C., Yeatman, E., Banavar, M., Constantinides, A.G., Willerton, M., Spanias, A., Thornton, T., Manikas, A.: Performance comparison of localization techniques for sequential WSN discovery. In: Sensor Signal Processing for Defense (SSPD 2012), pp. 33–33 (2012)
4. Bachrach, J., Taylor, C.: Localization in sensor networks. In: Handbook of Sensor Networks: Algorithms and Architectures, pp. 277–310 (2005)
5. Farina, M., Deb, K., Amato, P.: Dynamic multiobjective optimization problems: test cases, approximations, and applications. IEEE Trans. Evol. Comput. **8**, 425–442 (2004)
6. Zhou, A., Qu, B.-Y., Li, H., Zhao, S.-Z., Suganthan, P.N., Zhang, Q.: Multiobjective evolutionary algorithms: a survey of the state of the art. Swarm Evol. Comput. **1**, 32–49 (2011)
7. Mao, G., Fidan, B., Anderson, B.D.O.: Wireless sensor network localization techniques. Comput. Netw. **51**, 2529–2553 (2007)
8. Mirjalili, S., Mirjalili, S.M., Lewis, A.: Grey Wolf Optimizer. Adv. Eng. Softw. **69**, 46–61 (2014)
9. Zavala, G.R., Nebro, A.J., Luna, F., Coello Coello, C.A.: A survey of multiobjective metaheuristics applied to structural optimization (2014)
10. Mirjalili, S., Saremi, S., Mirjalili, S.M., Coelho, L.D.S.: Multi-objective grey wolf optimizer: a novel algorithm for multi-criterion optimization. Expert Syst. Appl. **47**, 106–119 (2016)
11. Knowles, J.D., Corne, D.W.: Approximating the nondominated front using the Pareto archived evolution strategy. Evol. Comput. **8**, 149–172 (2000)

Evaluation of Mobile Phone Wireless Charging System Using Solar and Inductive Coupling

Norizam Sulaiman(✉), Mohd Shawal Jadin, Amran Abdul Hadi, Muhamad Sharfi Najib, Mohd Ashraf Ahmad, and Nurul Shazrien Amsyha Yusuf

Faculty of Electrical and Electronics Engineering, Universiti Malaysia Pahang, 26600 Pekan, Pahang, Malaysia
{norizam,mohdshawal,amranhadi,sharfi,mashraf}@ump.edu.my, shazrienamysha.sa@gmail.com

Abstract. The wireless charging system now becomes one of the emerging technologies especially in the application of communication systems and beneficial to the wireless electronic appliances. Among them are mobile phones, cameras, personal digital assistance (PDA), cooler, torchlight and drill. Those wireless devices require battery to store and provide power before the device can be used. Hence, in order to solve the problem of short life of the battery of mobile phone, this project proposes adding a solar charging system base on inductive coupling method to the mobile phone to improve the usage of mobile phone in term of standby time, talk-time, online applications and power consumption especially in the remote area. Inductive coupling is among the effective method in wireless charging system to charge electronics device and reduce the constraint of the power cord or wired system. Meanwhile, solar cell is among the energy harvesting devices that is widely employed in many electronics application. The outcome of the project describes the comparison of the power consumption between the wire charging systems with solar-based wireless charging system. From the analysis of the results, solar-powered mobile phone with inductive coupling produced 21 h 46 min standby time after charging for 13 h 15 min compared to the existing charging system (wired system) which produce 17 h 5 min standby time after charging for 2 h 30 min. In addition, proposed system has high power consumption in term of standby time, talk-time and online application. Based on the results of the project, it could suggest that the wireless solar-powered mobile phone can replace the existing charging system in term of standby time.

Keywords: Wireless charging system · Mobile phone · Solar · Inductive coupling · Standby time

1 Introduction

Nowadays, the portable equipment that depends on the solar energy as a power supply is generally utilized as a part of human daily life [1]. With the expansion of the energy demand and the concern of ecological contamination around the

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_26

globe, the usage of photovoltaic (PV) system is preferable and become eminent system [1]. Due to the energy crisis, renewable energy sources such as solar become top energy sources to provide power to portable devices.

The solar energy which specifically generated by the sun, can be converted to electrical energy by utilizing PV cell. For example, one of the applications that utilizing PV innovation is wireless battery charger to extend the battery life of the electronic gadgets without the help of wires [2,3].

Despite of the drawbacks of the application of wireless system in electronic devices such as slower charging time and radiation effect to human body, the technology of the wireless system keep evolving and widely used in many electronic appliances and reduce the wired-based appliances especially in mobile phone charging system. The wireless charging system or wireless power transmission has been employed for couple of decades. Researchers have come out with several techniques in wireless charging system in order to find the effective method. One of them is inductive coupling which was first implemented by researcher in 2007 and was effective in transferring power wirelessly [4].

This project concentrates on utilizing solar cell energy as a part of giving an option in selecting the power source to charge the mobile phones and electronic gadgets. Meanwhile, the inductive coupling method is employed as a wireless communication medium to transfer power from the source (solar) to the load (mobile phone). Here, there are many previous works and researches that related to this field. For example, the solar harvesting system that expanded as a top of the line implanted system was depicted in which a switch network was utilized to provide electrical energy to an individual system segments either from the sun powered board or the battery [5–7]. Aside from sun oriented energy, this project additionally exhibits a wireless system for power charging by transferring the solar energy through the wireless signal in which the mobile phone can be charged without a wire connection. Next, the performance of the solar-powered mobile phone (wireless charging system) is measured in term of standby time and talk-time with existing mobile phone charging system and power consumption for both systems.

2 Literature Review

Nowadays, energy consumption becomes crucial due to the increasing world population and applications of the device. Researchers or scientists have invented various energy harvesting device to produce renewable energy from environment such as solar, wind, tidal wave, thermal, radio frequency and vibration. Among them, the solar and wind are the famous renewable and non-contaminating source of energy to run many application of the devices. For example, solar energy is used to provide sufficient power or electrical energy to run the mobile phones, calculators, PDAs, wristwatches, torchlight, fan, laptop, television, fluorescent lighting, wireless mouse, Bluetooth headset, MP3 player, drill and cooler [7,8]. However, solar is more preferable for energy harvester compared to wind since wind turbine is required to produce energy from the wind which more expensive and costly compared to solar.

Table 1. Previous researches on wireless charging techniques

Technique	Advantage	Disadvantage	Charging distance	Application
Inductive coupling	Safe for human, simple implementation	Short charging distance warming impact, not appropriate for vehicle applications	From a few mm to a few cm	Smart phones, toothbrush cameras, RFID tags, smart cards
Magnetic resonance coupling	Free arrangement amongst chargers and charging various devices all the while on various power, high charging proficiency, non-viewable pathway charging	Not appropriate for portable applications, restricted charging distance and complex execution	From a few cm to a few m	smart phone, home appliances (TV and desktop), electric vehicle charging
Microwave radiation	Long viable charging distance, appropriate for portable applications	Not protected when the RF density exposure is high, low charging effectiveness, viewable pathway charging	Typically within several tens of meters, up to several kilometres	LEDs, RFID cards, wireless sensors and implanted body devices

Solar energy is produced from the sunlight using PV or solar cells after exposing the PV cells to the sunlight for certain time [9]. The PV cell then will convert the solar energy from the sunlight into electrical energy or direct current (DC). Here, solar charging controller is required before storing the solar energy into battery. The controller circuit will regulate the flow of charge (current) from the PV cells to the battery [9]. The design of the solar controller circuit is very crucial in order to provide sufficient current to run the electronic gadgets.

Lately, the emerging of the technology of the wireless systems has made it become the vital type of communication system and practical system especially in remote area. The popular method of wireless systems are cellular broadcast, wireless fidelity (Wi-Fi) and radio frequency waves (RF). The system provides the great contribution to humankind in application of electronic gadgets and devices. The wireless system via inductive coupling is quite new. However, the previous researchers have claimed that this method is more effective, efficient and practical in transferring energy across the free air or air gap [5,10–12]. The inductive coupling employs the magnetic field induction that conveys electrical energy between two coils from the source to the load. Here, the performance of the transfer of the energy is extremely depends on the tightness of the coupling of the coils and the quality of the components of the coils [11,12]. In wireless charging techniques, besides inductive coupling, magnetic resonance coupling and microwave radiation were also employed by researchers. The advantages and disadvantages among the techniques are elucidated by Table 1.

3 Methodology

This project is implemented to study the performance of the solar-based wireless charging system based on duration of charging time, standby time, talk-time and other mobile phone applications. The performance comparison is made between solar-based wirelesses charging system with power line charging system including charging system by power bank. The flow chart of project is depicted by Fig. 1.

In this project, the PV cell with dimension of 12 cm × 7.5 cm with peak power of 1 W and maximum voltage of 12 V is used to generate the electrical current (83.3 mA). Here, since the project has employed the small size of wireless charging system with 1 W and 12 V of PV cell, the PV cell is required to be exposed to a direct sunlight for duration more than 12 h. Meanwhile, the battery of the power bank is used to store the voltage and current received from solar cell. In order to boost up voltage and current of the wireless charging system, the solar controller circuit is designed to produce output voltage greater than 12 V using combination of voltage regulator, LM317, Zener diode, diode and resistor as shown in Fig. 2. The mobile phone with battery capacity of 1400 mAh is selected for this project. Besides, the Proteus software is employed to simulate the solar controller circuit in term of output voltage and current.

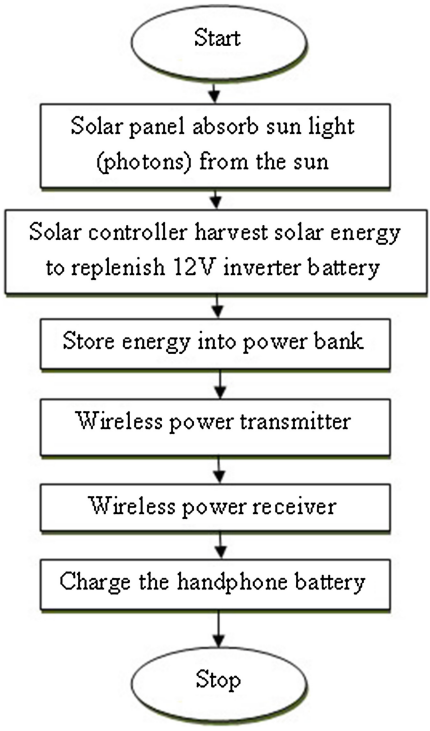


Fig. 1. Process flow of solar-powered wireless charging system.

For the transmission system in wireless charging system, the transmitter circuit consists of DC source, oscillator circuit and transmission coil. Here, the oscillator circuit consists of n-channel MOSFET IRF540 and diodes (IN4148). The value of the tank circuit of capacitor is 6.8 nF and the value of transmission coil is 0.67 H. Meanwhile, the receiver circuit consists of receiver coil, rectifier circuit (1N4007 diode and capacitor of 6.8 nF). The output of receiver circuit will be connected to mobile phone through USB cable to charge the battery of mobile phone. The transmitter and receiver circuit is illustrated by Figs. 3 and 4 respectively. Meanwhile, the overall prototype of the project is elucidated by Fig. 5 which consist circuits of transmitter with coil, receiver with coil, solar controller, and PV cell. The USB cable and mobile phone are also part of the prototype. The coil of transmitter and receiver is required to be located close to each other. Otherwise, they might disturb the performance of the power transfer from solar cell to mobile phone. The value of transmitter and receiver coil is measured using LCR meter. Again, the Proteus software is used to simulate the output parameters of the transmitter and receiver circuit.

4 Results and Discussion

The overall prototype of the circuit is constructed and tested successfully. Several parameters are required to be calculated before analyzing the performance of the wireless charging compared to wired-base charging.

4.1 Current Capacity of Solar Cell

The current from the solar cell is calculated using formula stated in (1). Thus, the solar panel with power of 1 W and maximum voltage of 12 V will produce current of 83.33 mAh. Thus, for the 1400 mAh battery of mobile phone with selected size of solar cell, it requires 16 h and 8 min to charge the mobile phone as described by (2).

$$I = \frac{P}{V} = \frac{1\text{W}}{12\text{V}} = 83.33\,\text{mA}. \tag{1}$$

$$\textit{Charging duration} = \frac{1400\,\text{mAh}}{83.33\,\text{mAh}} = 16\,\text{h}\ 8\,\text{min}. \tag{2}$$

4.2 Transmitter and Receiver Parameters

The working frequency of the transmitter and receiver is calculated using the following formula as shown in (3). The results are shown in (4) and (5).

$$F = \frac{1}{2\pi\sqrt{LC}}, \tag{3}$$

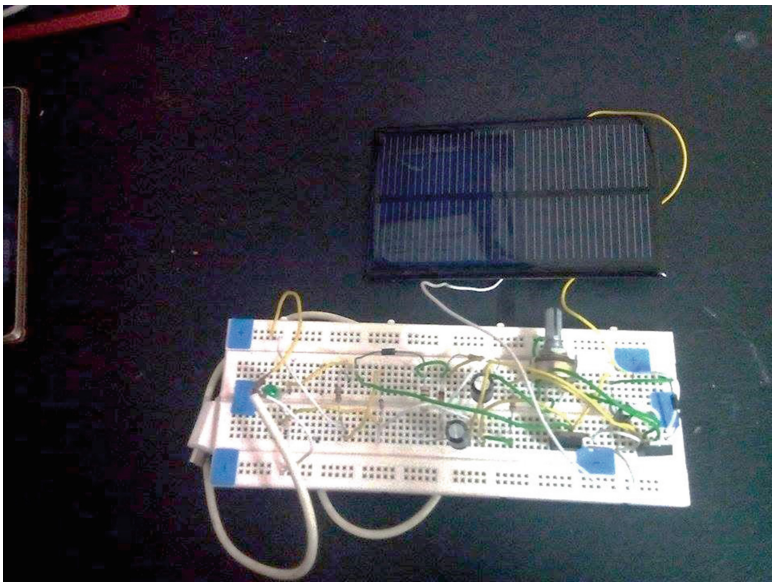


Fig. 2. The prototype of solar cell and controller circuit.

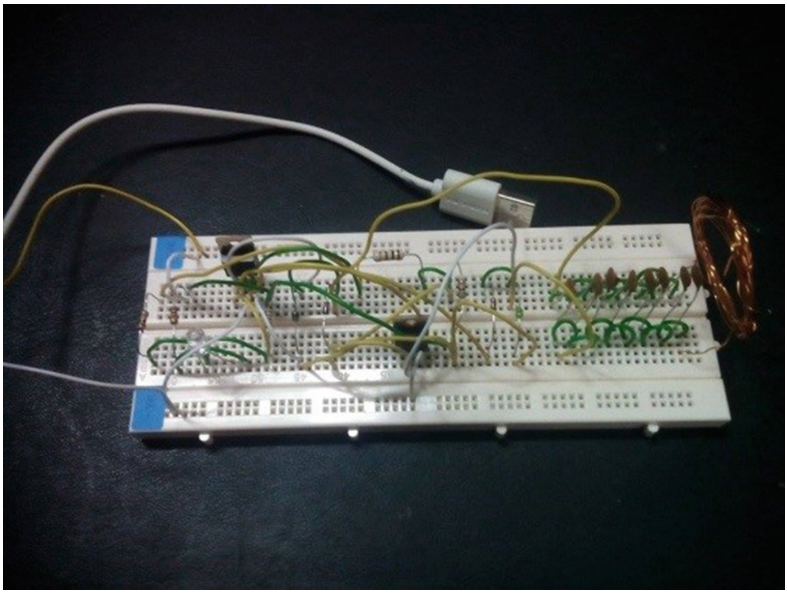


Fig. 3. The prototype of transmitter circuit.

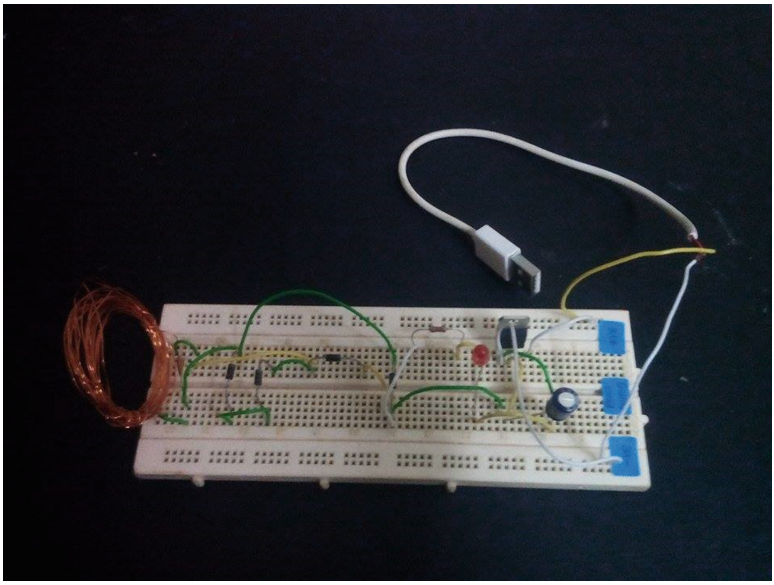


Fig. 4. The prototype of receiver circuit.

$$F = \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2\pi\sqrt{(6.8 \times 10^{-9})(0.674 \times 10^{-6})}} = 2.35 \times 10^{6}\,\text{Hz}, \quad (4)$$

$$F = \frac{1}{2\pi\sqrt{LC}} = \frac{1}{2\pi\sqrt{(6.8 \times 10^{-9})(1.235 \times 10^{-6})}} = 1.74 \times 10^{6}\,\text{Hz}, \quad (5)$$

For the transmitter part, the tank circuit will produce 6.8 nF capacitor and inductance of transmission coil of 0.674 H. By applying these values using formula in equation will produce transmission frequency of 2.35 MHz. Meanwhile, for the receiver parameter, by applying the value of receiver coil of 1.235 H and capacitor value of 6.8 nF, the receiver will run at the frequency of 1.74 MHz.

The charging performance between power line system (wired-base) and wireless system is performed. In order to come out with the solid results, the same mobile phone with battery capacity of 1400 mAh is used. The results of the wired system and wireless charging system are depicted by Table 2. Apparently, wireless system outperforms wired system in standby time but required longer charging time.

Finally, the comparison of the power consumption between wireless charging system and wire-based (power line) charging system is done and shown in Fig. 6. Based on Fig. 6, it is apparent that power consumption of the wireless charging is greater than wire-based charging system in term of standby time, talk-time,

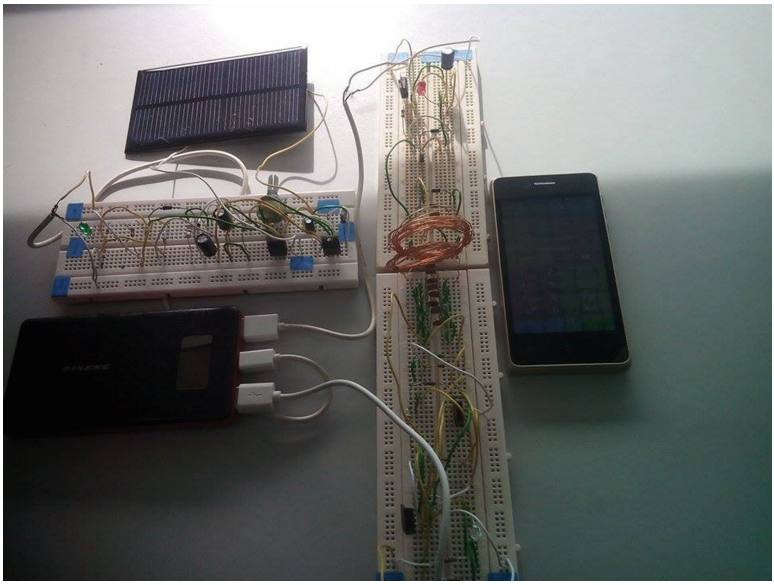


Fig. 5. The overall prototype of the circuit.

active online usage, and video playback. However, the wireless charging system requires longer charging time compared to wired system.

The simulation process is also carried out for the circuitry of the wireless system in term of voltage at controller, voltage at transmitter, voltage at receiver, frequency at transmitter and frequency at receiver as shown in Table 3. Based on Table 3, it can be seen that the inductive coupling produced higher current at the receiver part. The voltage of 3.56 V that received by receiver is good since it just enough to charge the Lithium-Ion battery of mobile phone. The simulation results also indicates that the solar controller circuit capable to provide sufficient voltage and current for solar-based wireless charging system.

Table 2. The Performance of the Wired System and Wireless Charging System

Parameters	Duration	
	(Wired system)	Solar wireless system
Charging time	2 h 30 min	13 h 15 min
Standby time	17 h 5 min	21 h 28 min
Talk-time/call time	6 h 57 min	4 h 13 min
Active online usage	7 h 10 min	5 h 51 min
Video playback time	7 h 5 min	3 h 1 min
Music playback time	11 h 52 min	11 h 52 min

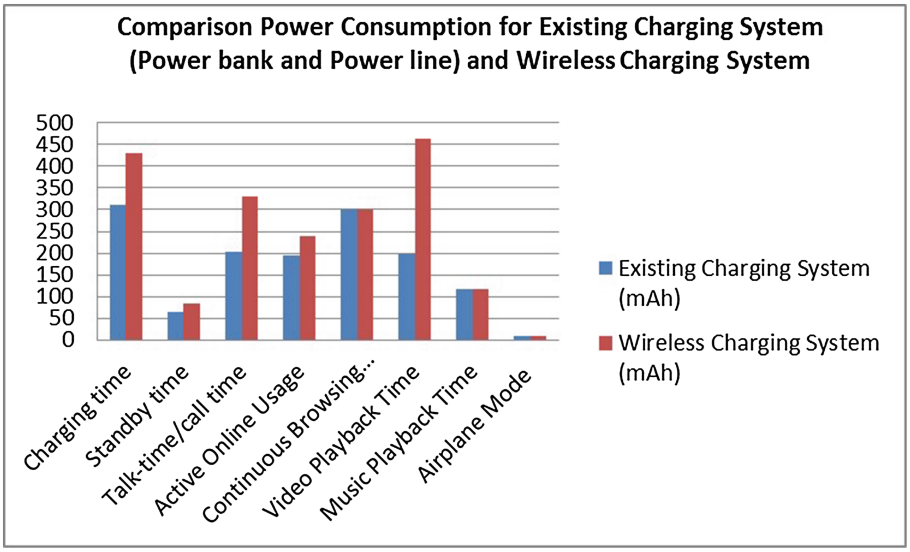


Fig. 6. The performance of power consumption.

Table 3. The simulation results of the wireless charging system

Parameters	Simulation results
Input voltage	12 V
Voltage at controller	10.96 V
Voltage at transmitter	8.2 V
Voltage at receiver	3.56 V
Current at transmitter	24.7 mA
Current at receiver	49.4 mA
Frequency at transmitter	2.35 MHz
Frequency at receiver	1.74 MHz

5 Conclusion

This study is performed to analyze the performance of the solar-based wireless charging system and wire-based charging system. The results of the study show that the solar-based wireless system based on inductive coupling is very good in term of standby time. The only disadvantage of the solar-based wireless system is longer charging duration but not so critical since the system can be used anywhere even in the remote area as long as the source of energy (sunlight) is there.

Acknowledgments. The author would like to thank and grateful to the undergraduate student and research team members, for their great support to make the project success. Also counted is Universiti Malaysia Pahang for providing financial support through research grant, RDU140376.

References

1. Ali, Q.I.: Design and implementation of a mobile phone charging system based on solar energy harvesting. In: 1st International Conference on Energy, Power, and Control, pp. 264–267. IEEE Press, New York (2010)
2. Singh, T., Bhardwaj, P., Singh, B., Kumar, A.: Design and development of portable power charger: a green energy initiative. Int. J. Electron. Commun. Technol. **5**, 1–2 (2014)
3. Anand, M., Kannan, Y.R.: Wireless power transfer by incorporation of solar energy. Int. J. Recent Dev. Eng. Technol. **3**, 85–91 (2014)
4. Mung, D.L., Lwin, K.S., Tun, H.M.: Design and construction of wireless charging system using inductive coupling. Int. J. Sci. Technol. Res. **4**, 282–287 (2015)
5. Bilal, N., Jilani, A., Hamid, H., Inayat, A., Naeem, S., Nimra: Circuit-model based analysis of wireless energy transfer system using inductive coupling. J. Emerg. Trends Comput. Inf. Sci. **5**, 746–749 (2014)
6. Cheng, T.C., Chang, W.C., Yarn, K.F., Kuo, J.K.: On the improvement of solar cell based photovoltaic charger system. In: 8th International Conference on Solid State and Integrated Circuit Technology Proceedings, Shanghai, pp. 986–988 (2006)
7. Tesla, N.: The transmission of electrical energy without wires as a means for further peace. Electr. World Eng. (1905)
8. Xiao, L., Niyato, D., Wang, P., Kim, D.I.: Wireless charger networking for mobile devices: fundamentals, standards and applications. J. IEEE Wireless Commun. (2015)
9. Kuo, J.K.: On the improvement of solar cell based photovoltaic charger system. In: International Conference on Solid-State and Integrated Circuit Technology (2006)
10. Zhu, B., Li, J., Hu, W., Gao, X.: Review of magnetic coupling resonance wireless energy transmission. Int. J. U and E Ser. Sci. Technol. (2015)
11. Costanzo, A.: Electromagnetic energy harvesting and wireless power transmission: a unified approach. IEEE J. **102**, 1692–1711 (2009)
12. Madawala, U.K., Thrimawithana, D.J.: Modular based inductive power transfers system for high power application. J. IET Power Electron. **5**, 1119–1126 (2007)

Hardware Implementation of MFCC Feature Extraction for Speech Recognition on FPGA

Van-Lan Dao(✉), Van-Danh Nguyen, Hai-Duong Nguyen, and Van-Phuc Hoang

Faculty of Radio-Electronics, Le Quy Don Technical University, Hanoi, Vietnam
kqha1025@gmail.com, vandanhktqs@gmail.com, mta.haiduongnguyen@gmail.com, phuchv@mta.edu.vn
http://www.lqdtu.edu.vn

Abstract. In this paper, an FPGA-based Mel Frequency Cepstral Coefficient (MFCC) IP core for speech recognition is presented. The implementation results on FPGA show that the proposed MFCC core achieves higher resource usage efficiency compared with other designs.

1 Introduction

Nowadays, speech recognition and speaker recognition researched and applied a lot of human life [1,2,4–7]. Feature extraction of speech is one of the most important issues in the field of speech recognition or speaker recognition. This phase is proceeded right after the input speech is pre-emphasized and windowed. There are two dominant acoustic measurements of speech signal. One is the parametric approach, which is developed to match closely the resonant structure of the human vocal tract that produces the corresponding speech sound. It is mainly derived from linear predictive analysis, such as LPC-based cepstrum (LPCC). The other one is the non-parametric method modeling the human auditory perception system. Mel frequency cepstral coefficients (MFCCs) are utilized for this purpose [8]. In recent studies of speech recognition system, the MFCC parameters perform better than others in the recognition accuracy [9,10]. Figure 1 presents the general model of a speech/speaker recognition. The objective of this paper is to design a low area MFCC core on FPGA with an improved MFCC algorithm for speech recognition and speaker recognition applications.

The rest of this paper is organized as follows. Section 2 describes the MFCC algorithm and Sect. 3 shows the MFCC core design. Section 4 presents the implementation results and finally, Sect. 5 concludes of the paper.

2 MFCC Algorithm

The extraction of the best parametric representation of acoustic signals is an important task to produce a good recognition performance. The efficiency of this phase is important for the next phase since it affects its behavior.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_27

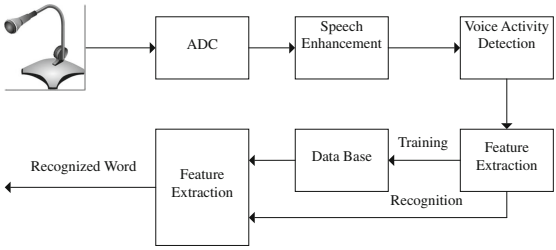


Fig. 1. General model of a speech/speaker recognition.

MFCC is based on the human hearing perception which cannot perceive frequencies over 1 KHz. In other words, MFCC is based on known variation of the human ear's critical bandwidth with frequency [8–10]. MFCC has two types of filter which are spaced linearly at low frequency below 1000 Hz and logarithmic spacing above 1000 Hz. A subjective pitch is present on Mel frequency scale to capture the important phonetic characteristic in speech.

The overall block diagram for MFCC algorithm is shown in Fig. 2 [11].

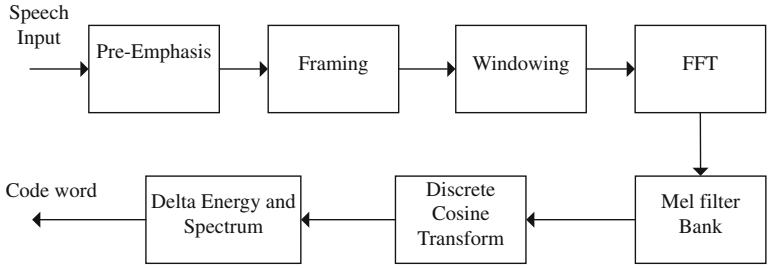


Fig. 2. MFCC block diagram.

MFCC consists of seven computational steps as shown in Fig. 2. The mathematical expression of each step is discussed briefly as follow.

Step 1: Pre-emphasis

This step processes the passing of signal through a filter which emphasizes higher frequencies. This process will increase the energy of signal at higher frequency.

$$Y[n] = X[n] - aX[n-1] \tag{1}$$

Lets consider $a = 0.95$, which make 95 % of any one sample is presumed to originate from the previous sample.

Step 2: Framing

The process of segmenting the speech samples obtained from analog to digital conversion (ADC) into a small frame with the length within the range of 20 ms to

40 ms. The voice signal is divided into frames of N samples. Adjacent frames are being separated by $M(M < N)$. Typical values used are $M = 100$ and $N = 256$.

Step 3: Hamming windowing
Hamming window is used as window shape by considering the next block in feature extraction processing chain and integrates all the closest frequency lines. The Hamming window equation is given as:

If the window is defined as $W(n), 0 \leq n \leq N-1$ where
$N =$ number of samples in each frame
$Y(n) =$ Output signal
$X(n) =$ input signal $W(n) =$ Hamming window, then the result of windowing signal is shown below:

$$Y[n] = X[n].W[n] \tag{2}$$

$$W(n) = 0.54 - 0.46 * \cos(\frac{2\pi n}{N-1}); 0 \leq n \leq N-1; \tag{3}$$

Step 4: Fast Fourier Transform (FFT)
FFT converts each frame of N samples from time domain into frequency domain. The Fourier Transform is to convert the convolution of the glottal pulse $U[n]$ and the vocal tract impulse response $H[n]$ in the time domain. This statement supports the equation below:

$$Y(w) = FFT[h(t) * X(t)] = H(w) * X(w) \tag{4}$$

If $X(w)$, $H(w)$ and $Y(w)$ are the Fourier Transform of $X(t)$ and $Y(t)$ respectively.

Step 5: Mel Filter Bank Processing
The frequency range in FFT spectrum is very wide and voice signal does not follow the linear scale. The bank of filters according to Mel scale as shown in Fig. 3 is then performed.

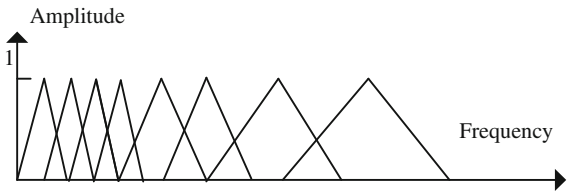


Fig. 3. Mel scale filter bank.

This figure shows a set of triangular filters that are used to compute a weighted sum of filter spectral components so that the output of process approximates to a Mel scale. Each filter's magnitude frequency response is triangular

in shape and equal to unity at the centre frequency and decrease linearly to zero at centre frequency of two adjacent filters [13]. Then, each filter output is the sum of its filtered spectral components. After that the following equation is used to compute the Mel for given frequency f in Hz:

$$Mel(f) = 2595 * \log 10(1 + {}^{f}/_{700}) \tag{5}$$

Step 6: Discrete Cosine Transform
This is the process to convert the log Mel spectrum into time domain using Discrete Cosine Transform (DCT). The result of the conversion is called Mel Frequency Cepstrum Coefficient. The set of coefficient is called acoustic vectors. Therefore, each input utterance is transformed into a sequence of acoustic vector.

Step 7: Delta Energy and Delta Spectrum
The voice signal and the frames changes, such as the slope of a formant at its transitions. Therefore, there is a need to add features related to the change in cepstral features over time axis. 13 delta or velocity features (12 cepstral features plus energy), and 39 features a double delta or acceleration feature are added. The energy in a frame for a signal x in a window from time sample t_1 to time sample t_2, is represented by the below equation:

$$Energy = \sum_{t=t_1}^{t_2} X^2(t) \tag{6}$$

Each of the 13 delta features represents the change between frames in the Eq. (7) corresponding cepstral or energy feature, while each of the 39 double delta features represents the change between frames in the corresponding delta features.

$$d(t) = \frac{c(t+1) - c(t-1)}{2} \tag{7}$$

In [13], the authors presented an improved MFCC algorithm which will be used in this paper.

3 MFCC Hardware Core Design

MFCC architecture presents as shown in Fig. 4. In this work, the MFCC core is designed with 80-point sub-frame and 128-point FFT. Data input is speech signal samples saved in RAM. The controller block is a finite state machine including two states as show Fig. 5.

4 Implementation Results

The MFCC core was implemented with VHDL code and implemented in Xilinx Virtex-VI-LX240T and Spartan 3A DSP 1800 FPGA platforms. The FPGA implementation results are shown in Table 1 in which the MFCC core area can be reduced lower than other designs. Compared to other designs, this work can reduce the number of DSP48E1s significantly and numbers of slice LUTs/Registers are also lower than in [14–16].

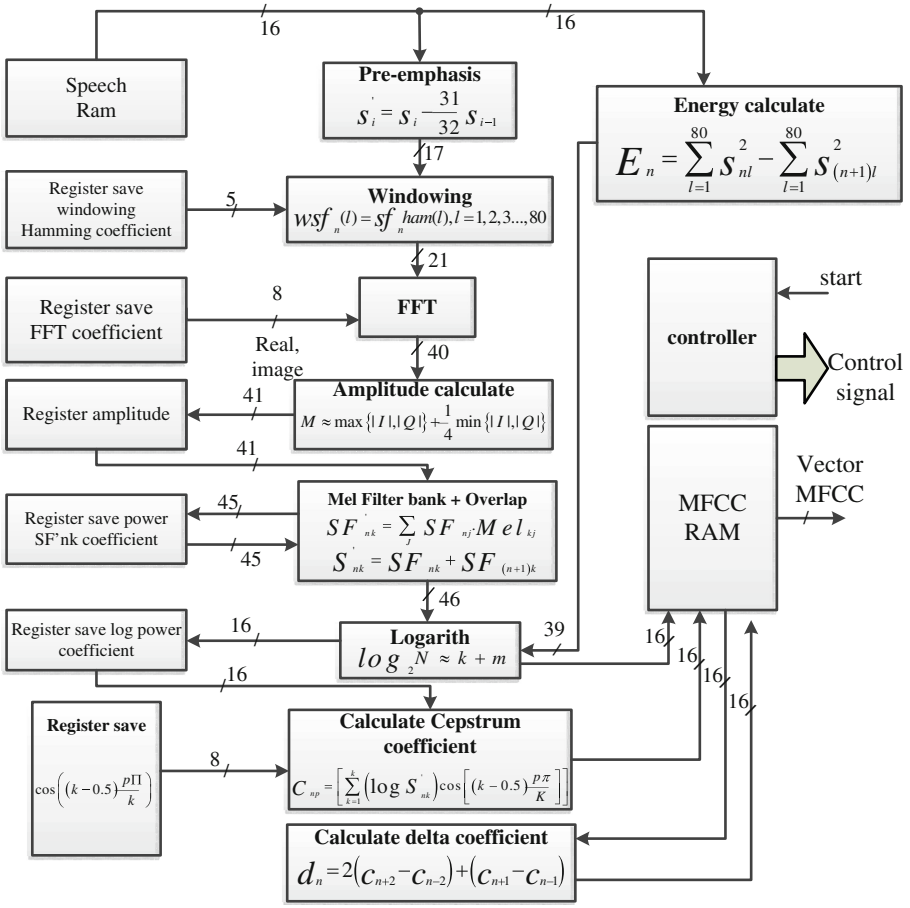


Fig. 4. Mel scale filter bank.

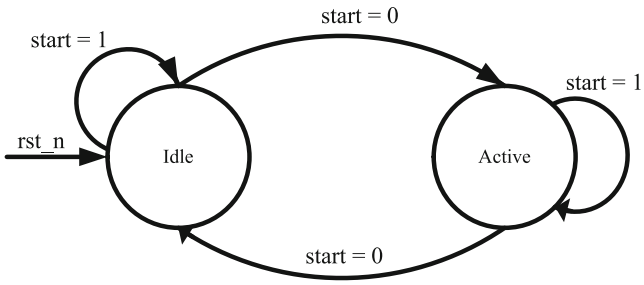


Fig. 5. Finite state machine for controller block.

Table 1. Implementation results of MMFC core on Xilinx FPGA.

FPGA device	Virtex 6-lx240T [14]	Virtex 6-lx240T [15]	Spartan 3A DSP 1800 [16]	Virtex 6-lx240T (Our design)	Spartan 3A 1800 (Our design)
Number of Slice Registers	11577	13726	23251	11567	11335
Number of Slice LUTs	8193	14837	35104	7353	11581
Number of fully used LUT-FF pairs	-	-	39452	1228	10700
Number of Block RAM/FIFO	44	4	14	8	15
Number of DSP48E1s	82	57	33	10	11
Frequency (MHz)	155	65.5	-	139	52

5 Conclusion

This paper has presented an area efficient MFCC core for speech/speaker recognition applications. The implementation results in FPGA show the MFCC core area can be reduced significantly. In the future, we will optimize the power consumption and area for the MFCC core and apply it for speech/speaker recognition applications.

References

1. Boves, L., den Os, E.: Speaker recognition in telecom applications. In: Proceedings of the Interactive Voice Technology for Telecommunications Applications, Torino, pp. 203–208 (1998)
2. McLoughlin, I.V., Sharifzadeh, H.R.: Speech recognition engine adaptions for smart home dialogues. In: Proceedings of the International Conference on Information, Communications and Signal Processing, Singapore, pp. 1–5 (2007)
3. Marchetto, E., Avanzini, F., Flego, F.: An automatic speaker recognition system for intelligence applications. In: Proceedings of the European Signal Processing, Glasgow, pp. 1612–1616 (2009)
4. Selvan, K., Joseph, A., Anish Babu, K.K.: Speaker recognition system for security applications. In: IEEE Recent Advances in Intelligent Computational Systems (RAICS), pp. 26–30 (2013)
5. Ajgou, R., Sbaa, S., Ghendir, S., Chamsa, A., Taleb-Ahmed, A.: Robust remote speaker recognition system based on AR-MFCC features and efficient speech activity detection algorithm. In: International Symposium on Wireless Communications Systems (ISWCS), Barcelona, pp. 722–727 (2014)
6. Malode, A.A., Sahare, S.L.: An improved speaker recognition by using VQ and HMM. In: Proceedings of the International on Sustainable Energy and Intelligent Systems (SEISCON 2012), Tiruchengode, pp. 1–7 (2012)
7. Lung, V.D., Truong, V.N.: Vietnamese speech recognition using dynamic time warping and coefficient of correlation. In: Proceedings of the International Conference on Control, Automation and Information Sciences (ICCAIS), Nha Trang, pp. 64–67 (2013)
8. Tuzun, O.B., Demirekler, M., Nakiboglu, K.B.: Comparison of parametric and non-parametric representations of speech for recognition. In: Proceedings, pp. 65–68 (1994)

9. Openshaw, J.P., Sun, Z.P., Mason, J.S.: A comparison of composite features under degraded speech in speaker recognition. In: Proceedings on Acoustics, Speech, and Signal Processing, vol. 2, Minneapolis, USA, pp. 371–374 (1993)
10. Vergin, R., O'Shaughnessy, D., Gupta, V.: Compensated mel frequency cepstrum coefficients. In: Proceedings on Acoustics, Speech, and Signal Processing, Minneapolis, USA, pp. 323–326 (1996)
11. Ibrahim, N.J., et al.: Quranic verse recitation feature extraction using Mel-frequency cepstral coefficients (MFCC). In: Proceedings of the International Colloquium on Signal Processing and Its Applications (CSPA), Kuala Lumpur, Malaysia (2008)
12. Price, J., Sophomore Student: Design an automatic speech recognition system using maltab. University of Maryland Estern Shore Princess Anne
13. Wang, J.-C., Wang, J.-F., Weng, Y.-S.: Chip design of mel frequency cepstral coefficients for speech recognition. In: Proceedings of the Advanced IEEE International Conference on Acoustics, Speech, and Signal Processing, Istanbul, vol. 6, pp. 3658–3661 (2000)
14. Wassi, G., Iloga, S., Romain, O., Granado, B.: FPGA-based real-time MFCC extraction for automatic audio indexing on FM broadcast data. In: Proceedings on Design and Architectures for Signal and Image Processing (DASIP), Krakow, pp. 1–6 (2015)
15. Bahoura, M., Ezzaidi, H.: Hardware implementation of MFCC feature extraction for respiratory sounds analysis. In: Proceedings of the International Workshop on Systems, Signal Processing and their Applications (WoSSPA), Algiers, pp. 226–229 (2013)
16. Ehkan, P., Zakaria, F.F., Warip, M.N.M., Sauli, Z., Elshaikh, M.: Hardware implementation of MFCC-based feature extraction for speaker recognition. In: Sulaiman, H.A., Othman, M.A., Othman, M.F.I., Rahim, Y.A., Pee, N.C. (eds.) Advanced Computer and Communication Engineering Technology. LNEE, vol. 315, pp. 471–480. Springer, Heidelberg (2015). doi:10.1007/978-3-319-07674-4_46

Improved Adaptive Fuzzy Sliding Mode Control for Second Order Nonlinear System

Pham Van Thiem[1(✉)], Lai Khac Lai[1], and Nguyen Thi Thanh Quynh[2]

[1] Automatic Control Department, Electronics Faculty,
Thai Nguyen University of Technology, Thai Nguyen City, Viet Nam
{phuthiem,laikhaclai}@tnut.edu.vn
[2] CReSTIC, Universite de Reims Champagne Ardenne, Reims, France
thi-thanh-quynh.nguyen@etudiant.uni-reims.fr
http://www.tnut.edu.vn

Abstract. In this paper, an Adaptive Fuzzy Sliding Mode combined with Model Reference Adaptive System (MRAS) is proposed to a single input-single output nonlinear system. The main goal of this paper is the control of second order nonlinear system. Firstly, adaptive fuzzy sliding mode is applied to design the controller. Next, a performance of system and a chattering phenomenon are improved by adjusting parameters of the controller via MRAS. Finally, an example is presented to illustrate the proposed methods. By applying Lyapunov stability theory, the adaptive law that is derived in this study is robust and convergent quickly. The simulation results and analysis show that the proposed method is better than Adaptive Fuzzy Sliding Mode in the sense of robustness against disturbance and reduce the chattering.

Keywords: Adaptive fuzzy sliding mode · Model reference adaptive systems (MRAS) · Coupled tank liquid level

1 Introduction

The nonlinear systems have attracted widespread attention in the recent decades [1,2,4,6]. There are many controlled methods derived from linear controller as gain scheduling, Jacobi matrix [5]. However, a question is that of when there exists a global change of coordinates, a diffeomorphism, in the state space that carries the nonlinear system into linear system. Krener [10] showed the importance of the Lie algebra of a vector fields associated with the system in studying such a question and gave an answer to this problem. D. Cheng [7] illustrated that we always find the coordinates transformation in order to transform a given nonlinear system with outputs into a controllable and observable linear one.

With the nonlinear system [3,5,6] K. Khalil use the exact feedback linearization approach to define a control law with assumption that the modeled object is clearly known and states are measured. If we do not know exactly parameters of object, two trends will appear. The first trend consists in introducing

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_28

a direct adaptive fuzzy control combined with sliding mode control [3,5] based on Lyapunov theory. The second trend, with which we are concerned, is carried out in controller design based on universal approximation [1,2,5], the indirect adaptive fuzzy control (IAFC). The research of it under sliding mode control (SMC) has attracted much attention. The apparent similarities between SMC and IAFC motivate considerable research efforts in combining the two achieving more superior performances such as overcoming some limitations of the traditional SMC. However, SMC suffers from a well-known problem chattering due to the high gain and high-speed switching control. The undesirable chattering may excite previously unmodelled system dynamics and damage actuators, resulting in unpredictable stability. On the other hand, it often remains the estimation error when we use universal approximation method via adaptive fuzzy system. The undesirable estimation error can make transient response longer, if it is large. Therefore, in this paper we consider this error in designing the adaptive fuzzy sliding mode controller.

Furthermore, the controller in [1–3,5] guarantees still the stable closed loop system, but it does not concentrate on the performance of closed loop system, for example: the settling time, overshoot and error between the output signal and reference. Therefore, in this paper, we focus on mitigating three problems: decrease influencing chattering phenomenon, reduce the error between reference trajectory and output trajectory, and raise the settling time of transient response.

The remainder of this paper is organized as follows. In Sect. 2, we provide the adaptive fuzzy sliding mode control. Section 3 then considers the improved adaptive fuzzy sliding mode control. Simulation results are given in Sects. 4 and 5 concludes this work.

2 Adaptive Fuzzy Sliding Mode Control (AFSC)

Without loss of generality, we consider the dynamic nonlinear has two state variables given a structure as:

$$\begin{cases} \dot{y} = \dot{x}_1 = x_2 \\ \ddot{y} = \dot{x}_2 = F(\boldsymbol{x}) + G(\boldsymbol{x})u \ + d(x); \|d(\boldsymbol{x})\| \leq \rho \\ y = x_1 \end{cases} \tag{1}$$

where $\boldsymbol{x} = (x_1 \ \ x_2)^T \in R^2$; $u, y \in R$ and $F(\boldsymbol{x}); G(\boldsymbol{x}) \in C^\infty$ are smooth functions, $d(\boldsymbol{x})$ is disturbance. The control target of system (1) is stable tracking a prior trajectory $y_m(t)$ called reference trajectory, as means $\lim\limits_{t\to\infty} e(t) = 0, |e(t)|$ is finite, where $e(t) = y_m(t) - y(t)$ is tracking error.

We assume that functions $F(\boldsymbol{x}), G(\boldsymbol{x})$ are unknown exactly. Consequently, we need to evaluate an online function $F(\boldsymbol{x}), G(\boldsymbol{x})$ based on adaptive fuzzy mode algorithm. In the [3], consider the MISO fuzzy logic control system has two inputs $\boldsymbol{x} = (x_1 \ x_2)^T \in \upsilon \in \mathbf{R}^2$ and an output $\boldsymbol{f} \in \upsilon$. This fuzzy logic system

with the center-average defuzzifier, product inference, and singleton fuzzifier is of the following form:

$$f(\boldsymbol{x}) = \frac{\sum_{l=1}^{M} f^{-l}\left(\prod_{j=1}^{2} \mu_{A_j^l}(x_j)\right)}{\sum_{l=1}^{M}\left(\prod_{j=1}^{2} \mu_{A_j^l}(x_j)\right)} = \boldsymbol{\theta}^T \boldsymbol{\xi}(\boldsymbol{x}) \tag{2}$$

where $\boldsymbol{\theta}^T = (f^{-1}, ..., f^{-M})$ is a parameter vector of the output membership functions, and $\boldsymbol{\xi}(\boldsymbol{x}) = (\xi^1(x), ..., \xi^M(x))^T$ is a regressive vector with the regression $\xi^l(x)$.

$$\widehat{F}(\boldsymbol{x}) = \boldsymbol{\theta}_F^T \boldsymbol{\xi}_F(\boldsymbol{x}); \ \widehat{G}(\boldsymbol{x}) = \boldsymbol{\theta}_G^T \boldsymbol{\xi}_G(\boldsymbol{x}) \tag{3}$$

where the parameter vector $\boldsymbol{\theta}_F^T; \boldsymbol{\theta}_G^T$ are updated online so that the approximate error between $F(\boldsymbol{x}); G(\boldsymbol{x})$ and $\widehat{F}(\boldsymbol{x}); \widehat{G}(\boldsymbol{x})$ is minimal. Define the optimal parameter vector:

$$\boldsymbol{\theta}_F^* = \arg\min_{\boldsymbol{\theta}_F}(\sup_x |\boldsymbol{\theta}_F^T \boldsymbol{\xi}_F(\boldsymbol{x}) - F(\boldsymbol{x})); \boldsymbol{\theta}_G^* = \arg\min_{\boldsymbol{\theta}_F}(\sup_x |\boldsymbol{\theta}_G^T \boldsymbol{\xi}_G(\boldsymbol{x}) - G(\boldsymbol{x})) \tag{4}$$

If $\boldsymbol{\theta}_F \to \boldsymbol{\theta}_F^*; \boldsymbol{\theta}_G \to \boldsymbol{\theta}_G^*$ then $\widehat{F}(\boldsymbol{x}) \to F(\boldsymbol{x})$, $\widehat{G}(\boldsymbol{x}) \to G(\boldsymbol{x})$ that means the fuzzy system (3) can approximate smooth nonlinear functions with the arbitrary small error if the number of fuzzy rules is large enough [7]. The $F(\boldsymbol{x}), G(\boldsymbol{x})$ can be expressed:

$$F(\boldsymbol{x}) = \boldsymbol{\theta}_F^{*T} \boldsymbol{\xi}_F(\boldsymbol{x}) + \boldsymbol{\delta}_F(\boldsymbol{x}); \ \ |\boldsymbol{\delta}_F(\boldsymbol{x})| < \boldsymbol{\eta}_F$$

$$G(\boldsymbol{x}) = \boldsymbol{\theta}_G^{*T} \boldsymbol{\xi}_G(\boldsymbol{x}) + \boldsymbol{\delta}_G(\boldsymbol{x}); \ \ |\boldsymbol{\delta}_G(x)| < \boldsymbol{\eta}_G$$

We use sliding surface:

$$s(e) = k_p e + k_d \dot{e} \tag{5}$$

where k_p and k_d are coefficients of Hurwitz polynomial. The system is controlled in such a way that the state always more towards the sliding surface and hits it. The sign of control value must change at the intersection between the state trajectory and sliding surface. This controller will be made stable closed loop system.

$$u = \frac{1}{\widehat{G}(\boldsymbol{x})}(-\widehat{F}(\boldsymbol{x}) + \ddot{y}_m + k_p e + k_d \dot{e} + \eta s) + \frac{1}{\widehat{G}(\boldsymbol{x})}\eta_s \text{sgn}(s) = u_c(t) + u_s(t) \tag{6}$$

From (7), we see that if the parameters of systems have change by the time, the functions $F(\boldsymbol{x}), G(\boldsymbol{x})$ will change, and the control signal will be suitable adjusted for alternative parameters of object.

Assumption 1: the function $G(\boldsymbol{x}), \hat{G}(\boldsymbol{x})$ are bounded:$0 < G_{lo} \leq G(\boldsymbol{x}), \hat{G}(\boldsymbol{x}) \leq G_{up} < \infty$

Theorem 1: Consider the nonlinear system (1) with Assumption 1, the state feedback controller (6) and $\widehat{F}(\boldsymbol{x}), \widehat{G}(\boldsymbol{x})$ in (3) are applied with the $\boldsymbol{\theta}_F^T$ and $\boldsymbol{\theta}_G^T$ updated by adaptive law $\dot{\boldsymbol{\theta}}_F = -Q_F^{-1}\boldsymbol{\xi}_F(\boldsymbol{x})s(e)$; $\dot{\boldsymbol{\theta}}_G = -Q_G^{-1}\boldsymbol{\xi}_G(\boldsymbol{x})s(e)u$. The closed loop system will be asymptotic stable with $\eta_s \geq (\lambda + \rho); \lambda = \boldsymbol{\delta}_F(\boldsymbol{x}) + \boldsymbol{\delta}_G(\boldsymbol{x})u_c; \|d(\boldsymbol{x})\| \leq \rho$.

Proof: see Appendix

The system (1), (6), with the sliding surface (5) always exits a chattering phenomenon occurred when $s(e) = 0$ but $e(t) \neq 0$, the system do not approach the steady state, and the $u(t)$ with $\text{sgn}(s)$ component will sharply switch to ensure $s(e) = 0$. The chattering change dramatically if η_s is larger, but we choose small η_s the transient response is longer. Moreover, we do not choose η_s larger than $\lambda + \rho$.

3 Improved Adaptive Fuzzy Sliding Mode Control (IAFSC)

As in Sect. 2, to minimize the chattering phenomenon, we reduce two λ, ρ parameters, with $\lambda = \boldsymbol{\delta}_F(x) + \boldsymbol{\delta}_G(x)u_c$ and ρ in (1). When the system is affected by disturbance $d(x)$ we need adaptive law to decrease influencing disturbance that means ρ will be dropped.

Furthermore, in formula (6), two parameters $k_p; k_d$ of the state feedback controller are chosen in (5) that they still guarantee the stable closed loop system. However, it does not concentrate on performance of closed loop system, the settling time, overshoot and error between the output signal and reference. Therefore, to improve performance of system, concretely, the decrease error between reference trajectory and output trajectory, and raise the settling time, two parameters $k_p; k_d$ will be adjusted by adaptive law based on model adaptive reference system

The controller (6) can be rewritten as:

$$u = \frac{1}{\widehat{G}(\boldsymbol{x})}\left(-\widehat{F}(\boldsymbol{x}) + \ddot{y}_m + \theta_k^T\xi_e + \eta s\right) + \frac{1}{\widehat{G}(\boldsymbol{x})}\eta_s \text{sgn}(s) = u_c(t) + u_s(t) \quad (7)$$

where $\boldsymbol{\theta}_k^T = (k_p \ k_d)$; $\boldsymbol{\xi}_e = (e \ \ \dot{e})^T$. The $\boldsymbol{\theta}_k^T$ will be updated by the adaptive law based on reference system in Theorem 2.

Theorem 2: Consider the nonlinear system (1) with assumption 1, the state feedback controller (8) and $\widehat{F}(\boldsymbol{x}), \widehat{G}(\boldsymbol{x})$ in (3) are applied with the $\boldsymbol{\theta}_G^T$; $\boldsymbol{\theta}_F^T$ and $\boldsymbol{\theta}_k^T$ updated by adaptive law:

$$\dot{\boldsymbol{\theta}}_F = -Q_F^{-1}\boldsymbol{\xi}_F(\boldsymbol{x})\varepsilon; \ \dot{\boldsymbol{\theta}}_G = -Q_G^{-1}\boldsymbol{\xi}_G(\boldsymbol{x})\varepsilon u; \ \dot{\boldsymbol{\theta}}_k = Q_k^{-1}\boldsymbol{\xi}_e\varepsilon$$

The closed loop system will be asymptotic stable with $\eta_s \geq (\lambda + \rho)$.

Proof: The second derivative of the output error between output and reference is expressed:

$$\begin{aligned}
&\ddot{y} = \ddot{y}_m + \boldsymbol{\theta}_k^T \xi_e + \eta s + G(\boldsymbol{x})u_s - (\widehat{F}(\boldsymbol{x}) - F(\boldsymbol{x})) - (\widehat{G}(\boldsymbol{x}) - G(\boldsymbol{x}))u_c + d(\boldsymbol{x}) \\
&\dot{x}_{2m} - \dot{x}_2 = (\boldsymbol{\theta}_F^T \boldsymbol{\xi}_F(\boldsymbol{x}) - \boldsymbol{\theta}_F^{*T} \boldsymbol{\xi}_F(\boldsymbol{x}) - \boldsymbol{\delta}_F(\boldsymbol{x})) - G(\boldsymbol{x})u_s \\
&+(\boldsymbol{\theta}_G^T \boldsymbol{\xi}_G(\boldsymbol{x}) - \boldsymbol{\theta}_G^{*T} \boldsymbol{\xi}_G(\boldsymbol{x}) - \boldsymbol{\delta}_G(\boldsymbol{x}))u_c - d(\boldsymbol{x}) - \boldsymbol{\theta}_k^T \boldsymbol{\xi}_e - \eta s \\
&\dot{\varepsilon} = \tilde{\boldsymbol{\theta}}_F^T \boldsymbol{\xi}_F(\boldsymbol{x}) + \tilde{\boldsymbol{\theta}}_G^T \boldsymbol{\xi}_G(\boldsymbol{x})u_c - \boldsymbol{\theta}_k^T \boldsymbol{\xi}_e - G(x)u_s - d(\boldsymbol{x}) - \Delta(\boldsymbol{x}) - \eta s
\end{aligned}$$

where $\tilde{\boldsymbol{\theta}}_F^T = \left(\boldsymbol{\theta}_F^T - \boldsymbol{\theta}_F^{*T}\right), \tilde{\boldsymbol{\theta}}_G^T = \left(\boldsymbol{\theta}_G^T - \boldsymbol{\theta}_G^{*T}\right)$, are the parameter error and $\Delta(\boldsymbol{x}) = \boldsymbol{\delta}_F(\boldsymbol{x}) - \boldsymbol{\delta}_G(\boldsymbol{x})u_c$. Consider the Lyapunov candidate function:

$$V = \frac{1}{2}\varepsilon^2 + \frac{1}{2}\tilde{\boldsymbol{\theta}}_F^T \boldsymbol{Q}_F \tilde{\boldsymbol{\theta}}_F^T + \frac{1}{2}\tilde{\boldsymbol{\theta}}_G^T \boldsymbol{Q}_G \tilde{\boldsymbol{\theta}}_G^T + \frac{1}{2}\boldsymbol{\theta}_k^T \boldsymbol{Q}_k \boldsymbol{\theta}_k \; ; \boldsymbol{Q}_F > 0; \; \boldsymbol{Q}_G > 0; \; \boldsymbol{Q}_k > 0$$

where $\boldsymbol{Q}_i \in \mathbf{R}^{d \times d} (d = \dim \boldsymbol{\theta}_i)$ is a positive definite matrix. Take the derivative of V with respect to time and notice that $\tilde{\boldsymbol{\theta}}_i = \dot{\boldsymbol{\theta}}_i$, we have:

$$\begin{aligned}
&\dot{V} = \varepsilon\dot{\varepsilon} + \tilde{\boldsymbol{\theta}}_F^T \boldsymbol{Q}_F \dot{\boldsymbol{\theta}}_F + \tilde{\boldsymbol{\theta}}_G^T \boldsymbol{Q}_G \dot{\boldsymbol{\theta}}_G + \boldsymbol{\theta}_k^T \boldsymbol{Q}_k \dot{\boldsymbol{\theta}}_k = \tilde{\boldsymbol{\theta}}_F^T \boldsymbol{Q}_F \dot{\boldsymbol{\theta}}_F + \tilde{\boldsymbol{\theta}}_G^T \boldsymbol{Q}_G \dot{\boldsymbol{\theta}}_G + \boldsymbol{\theta}_k^T \boldsymbol{Q}_k \dot{\boldsymbol{\theta}}_k \\
&+\varepsilon(\tilde{\boldsymbol{\theta}}_F^T \boldsymbol{\xi}_F(\boldsymbol{x}) + \tilde{\boldsymbol{\theta}}_G^T \boldsymbol{\xi}_G(\boldsymbol{x})u_c - \boldsymbol{\theta}_k^T \boldsymbol{\xi}_e - G(\boldsymbol{x})u_s - d(\boldsymbol{x}) - \Delta(\boldsymbol{x}) - \eta s) \\
&\dot{V} = \tilde{\boldsymbol{\theta}}_F^T (\varepsilon \boldsymbol{\xi}_F(\boldsymbol{x}) + \boldsymbol{Q}_F \dot{\boldsymbol{\theta}}_F) + \tilde{\boldsymbol{\theta}}_G^T (\varepsilon \boldsymbol{\xi}_G(\boldsymbol{x})u + \boldsymbol{Q}_G \dot{\boldsymbol{\theta}}_G) - \varepsilon G(\boldsymbol{x})u_s \\
&\quad +\boldsymbol{\theta}_k^T (\boldsymbol{Q}_k \dot{\boldsymbol{\theta}}_k - \varepsilon \boldsymbol{\xi}_e) - \varepsilon d(\boldsymbol{x}) - \varepsilon \Delta(\boldsymbol{x}) - \eta s \varepsilon \\
&\dot{V} \leq -\eta s\varepsilon - \varepsilon G(\boldsymbol{x})u_s + |\varepsilon| \, \|\Delta(\boldsymbol{x})\| + |\varepsilon| \, \|d(\boldsymbol{x})\| \leq -\eta s\varepsilon - \varepsilon G(\boldsymbol{x})u_s + |\varepsilon| \, (\lambda + \rho) \\
&\dot{V} \leq -\eta \, |s| \, |\varepsilon| - |\varepsilon| \, (\lambda + \rho) \left(\frac{G(\boldsymbol{x})}{G_{lo}} - 1\right) \leq 0 \\
&where : -\varepsilon G(\boldsymbol{x})u_s + |\varepsilon| \, (\lambda + \rho) \leq 0 \Leftrightarrow \eta_s \geq (\lambda + \rho)
\end{aligned}$$

Choose the parameters update law as follow

$$\begin{aligned}
&\varepsilon \boldsymbol{\xi}_F(\boldsymbol{x}) + \boldsymbol{Q}_F \dot{\boldsymbol{\theta}}_F = 0 \Rightarrow \dot{\boldsymbol{\theta}}_F = -\boldsymbol{Q}_F^{-1} \varepsilon \boldsymbol{\xi}_F(\boldsymbol{x}) \\
&\varepsilon \boldsymbol{\xi}_G(\boldsymbol{x})u + \boldsymbol{Q}_G \dot{\boldsymbol{\theta}}_G = 0 \Rightarrow \dot{\boldsymbol{\theta}}_G = -\boldsymbol{Q}_G^{-1} \varepsilon \boldsymbol{\xi}_G(x)u; \boldsymbol{Q}_k \dot{\boldsymbol{\theta}}_k - \varepsilon \boldsymbol{\xi}_e = 0 \Rightarrow \dot{\boldsymbol{\theta}}_k = \boldsymbol{Q}_k^{-1} \varepsilon \boldsymbol{\xi}_e
\end{aligned} \tag{8}$$

where a weighting matrix Q_k given as:

$$\boldsymbol{Q}_k = \begin{pmatrix} q_{k11} & 0 \\ 0 & q_{k22} \end{pmatrix}; \dot{\boldsymbol{\theta}}_k = \boldsymbol{Q}_k^{-1} \boldsymbol{\xi}_e \varepsilon \Leftrightarrow \begin{pmatrix} \dot{k}_p \\ \dot{k}_d \end{pmatrix} = \begin{pmatrix} 1/q_{k11} & 0 \\ 0 & 1/q_{k22} \end{pmatrix} \begin{pmatrix} e \\ \dot{e} \end{pmatrix} \varepsilon$$

$$\begin{aligned}
k_d(t) &= k_d(0) + \tfrac{1}{q_{k22}} \textstyle\int (\dot{y}_m - \dot{y}) \varepsilon dt \\
k_p(t) &= k_p(0) + \tfrac{1}{q_{k11}} \textstyle\int (y_m - y) \varepsilon dt; \quad \varepsilon = \dot{y}_m - \dot{y}
\end{aligned} \tag{9}$$

and $k_p(0), k_d(0)$ are chosen in (5) to guarantee the stability of the closed loop system.

From (9), we can confirm that the affect of the disturbance $d(\boldsymbol{x}, t)$ will go down dramatically or eliminated by the adaptive laws (9). In other words, the normal of the disturbance ρ will decrease to zero. This leads to the normal of sum $\lambda + \rho$ equal to λ that mean the chattering phenomenon decrease significantly.

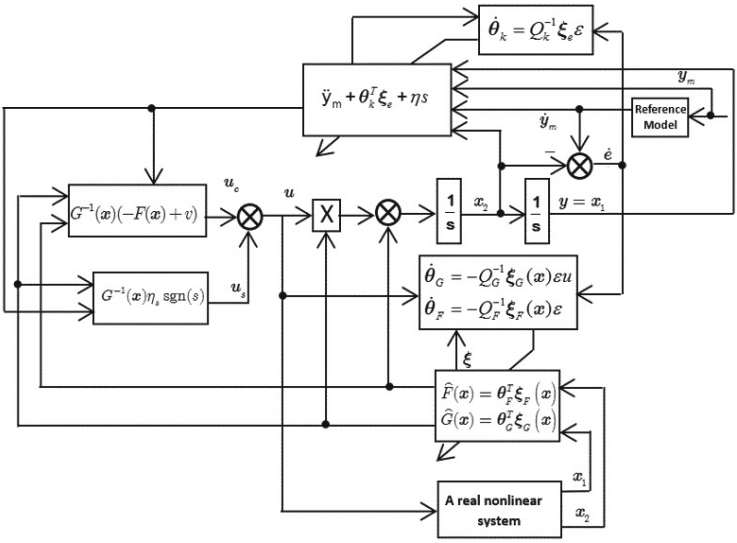


Fig. 1. The overall scheme improved adaptive fuzzy sliding mode control (IAFSC)

Reference Model: Explicit output reference, derivative output, and acceleration profile set point signals are created using the reference model, which is described by the transfer function:

$$H_{ref} = \frac{w_n^2}{s^2 + 2zw_n s + w_n^2} \tag{10}$$

The parameters of the reference model are chosen such as the higher order dynamics of the system will not be excited [9]. The following steps to design the controller via improved adaptive fuzzy sliding mode control system:

Step 1: The regressive vector $\boldsymbol{\xi}^l(\boldsymbol{x})$
Step 2: The adaptive law (8)
Step 3: Estimation the function: $\widehat{F}(\boldsymbol{x}) = \boldsymbol{\theta}_F^T\boldsymbol{\xi}_F(\boldsymbol{x})$; $\widehat{G}(\boldsymbol{x}) = \boldsymbol{\theta}_G^T\boldsymbol{\xi}_G(\boldsymbol{x})$
Step 4: The parameters of controller are updated by adaptive law as (9)
Step 5: The state feedback controller as (7)

4 Illustrative Example

In this section, we present an examples of tracking reference set point with a coupled tank liquid level system. The simulation illustrates the convergence of error under our proposed AFSC and IAFSC approach. In [13], the modeled equation of coupled tank liquid level system and parameters given as:

$$\frac{d}{dt}\begin{pmatrix} x_1 \\ x_2 \end{pmatrix} = \underbrace{\begin{pmatrix} \frac{-b_{12}a_{12}\sqrt{2g|x_1-x_2|}}{A_1} \\ \frac{1}{A_2}\left(b_{12}a_{12}\sqrt{2g\,|x_1-x_2|} - b_2a_2\sqrt{2gx_2}\right) \end{pmatrix}}_{f(\boldsymbol{x})} + \underbrace{\begin{pmatrix} \frac{k}{A_1} \\ 0 \end{pmatrix}}_{h(\boldsymbol{x})} u(t) \tag{11}$$

$$y = x_2 = g(\boldsymbol{x}); \boldsymbol{x} = (x_1\ x_2)^T = (h_1\ h_2)^T$$

We use the Theorem [11] in order to convert (11) into (1) where:

$$F(\boldsymbol{x}) = L_f^2 g(\boldsymbol{x}); G(\boldsymbol{x}) = L_h L_f g(\boldsymbol{x}) \tag{12}$$

Assuming that the function $f(\boldsymbol{x})$ and $g(\boldsymbol{x})$ describing the system dynamics are unknown, so we employ adaptive fuzzy system to online estimation $F(\boldsymbol{x})$ and $G(\boldsymbol{x})$. The control target guarantees for liquid level of Tank 2 at the set point. Opened valve ratio of pump 1 is adjusted by control law, and pump 2 open arbitrary.

We will calculate the $\widehat{F}(\boldsymbol{x})$; $\widehat{G}(\boldsymbol{x})$ by the adaptive fuzzy system. The inputs membership function are shaped in Gaussian over the interval [0 100] and $\boldsymbol{\theta}^T$ is parameter vector of output membership function which is updated by adaptive law in Theorem 1 for AFSC and Theorem 2 for IAFSC as Fig. 4. We choose the fix coefficients of AFSC and initial parameters $k_p(0) = 1$; $k_d(0) = 5$ of IAFSC

In this coupled tank liquid level system, we consider that this system is affected by disturbance $d(\boldsymbol{x}, t)$ which has changing frequency by the time as Fig. 6. With $\widehat{F}(\boldsymbol{x})$; $\widehat{G}(\boldsymbol{x})$ both AFSC and IAFSC could be evaluated when there is disturbance affect on the system. Comparison of liquid level control simulation between AFSC and IAFSC is present. Response of both AFSC and IAFSC are almost same (see Fig. 6).

However, it is obviously that the tracking error for the IAFSC due to reference liquid level is a lot less than those by AFSC (see Fig. 3). The adaptive $\boldsymbol{\theta}_k^T = (k_p\ k_d)$ of the IAFSC automatically alter when disturbance influence on system (see Fig. 5).

Table 1. The parameters of coupled liquid level system

Parameter	Value
A1, A2 the cross-sectional area of tank 1 and tank 2	100 cm2
Qin the flow rate of liquid into tank 1	u(cm3/sec)
Qout the flow rate of liquid into tank 2	0.8 cm3/sec
h1, h2 the height of liquid in tank1 and tank 2	cm
b2 the cross sectional area of outlet pipe in tank 2	0.5 cm2
a12 the cross-sectional area interaction pipe	0.5 cm2
b12 the value ratio of interaction pipe	1.5
g acceleration of gravity	981 cm2/sec
a2 the value ratio of outlet pipe of tank 2	1.5

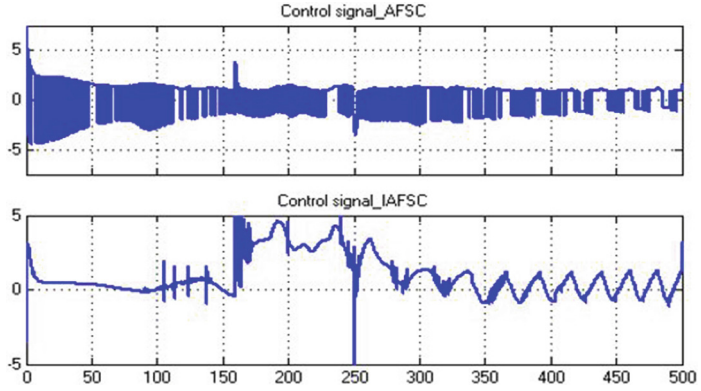


Fig. 2. Control signal with IAFSC ($\eta_s = 0.01$) and AFSC ($\eta_s = 0.4$)

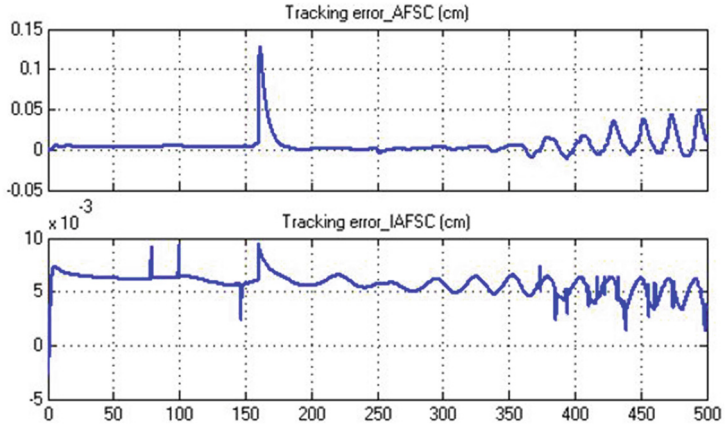


Fig. 3. Tracking error with AFSC (1st line) and with IAFSC (2nd line)

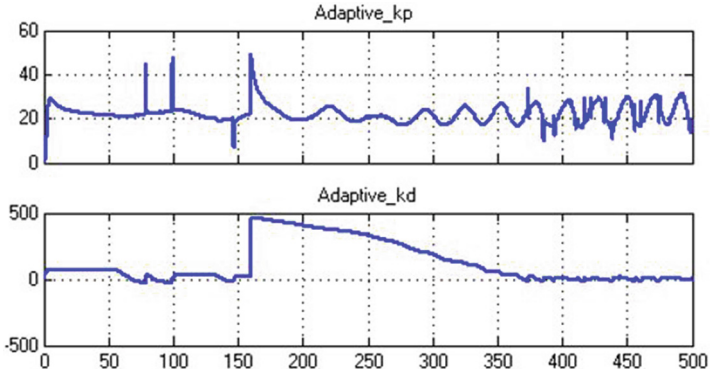


Fig. 4. Adaptive gains of the IAFSC

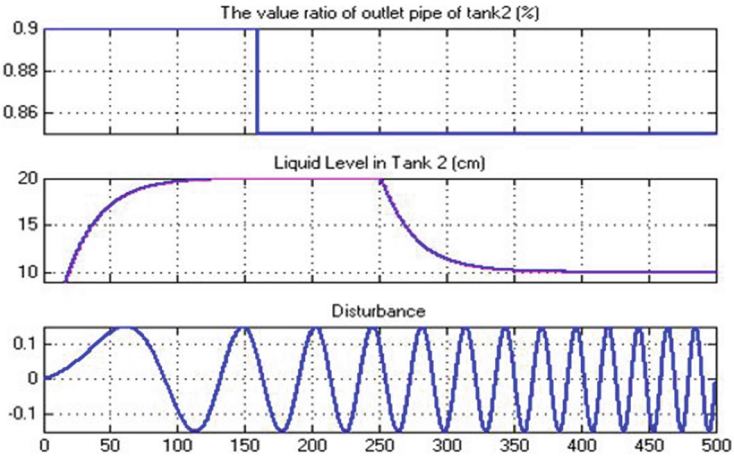


Fig. 5. Disturbance and tracking output with both of AFSC and IAFSC

The chattering change dramatically $\eta_s = 0.4$, but we choose small $\eta_s = 0.01$ combined with adaptive gains due to effect of disturbance. Not only the transient response is better, but also the chattering go down considerably (see Figs. 2, and 3)

5 Conclusion

In the paper, we have presented adaptive fuzzy sliding mode combined with MRAS, improved adaptive fuzzy sliding mode control offers a potential on deliver more accurate and high overall performance in the presence of all the preceding issues. We investigate the effect of the controller from the simulation results. Compared to the case with AFSC and IAFSC in the illustrated examples, coupled tank liquid level, for instance, can do the following (see Figs. 2, 3, 5, and 6): (a) Improve the transient behavior of the system; (b) Decrease the chattering to control signal; (c) Eliminate steady state errors; and (d) Decrease the influence of load disturbances and measurement noise. Strong properties achieved via the proposed method confirm that improved adaptive fuzzy sliding mode control is an attractive approach for controlling single input-single output nonlinear systems.

Acknowledgments. This research was supported by the MOET (The Ministry of Education and Training), Viet Nam, under grant no. B2016-TN01-01.

Appendix

The results of Theorem 1 are obtained in the same manner as Theorem 2.

References

1. Ren, X.M., Rad, A.B., Chan, P.T., Lo, W.L.: Identification and control of continuous-time nonlinear systems via dynamic neural networks. IEEE Trans. Ind. Electron. **50**(3), 478–486 (2003)
2. Bessa, W.M., Dutra, M.S., Kreuzer, E.: Adaptive fuzzy sliding mode control of underwater robotic vehicles. In: Proceedings of the XII International Symposium on Dynamic Problem of Mechanics, Brazil (2007)
3. Karsenti, L., Lamnabhi-Lagarrigue, F., Bastin, G.: Adaptive control of nonlinear parameterizaion. System Control Lett. **27**, 87–97 (1996)
4. Abid, H., Chtourou, M., Toumi, A.: Robust fuzzy sliding mode control for discrete nonlinear systems. Intl. J. Comput. Commun. Control. **III**(1), 6–20 (2008)
5. Wang, L.-X.: Stable adaptive fuzzy control of nonlinear system. IEEE Trans. Fuzzy Syst. **1**(2), 146–155 (1993)
6. Wang, L.-X.: Adaptive Fuzzy and Control. Prentice Hall, Upper Saddle River (1994)
7. Khalli, H.K.: Nonlinear Systems, 2nd edn. Prentice Hall, Upper Saddle River (1996)
8. Astolfi, A., Marconi, L.: Analysis and Design of Nonlinear Control Systems. Springer, Heidelberg (2008)
9. Phuoc, N.D.: Analysis and Control Nonlinear System. Science and Technics Publishing House, Hanoi (2012)
10. Krener, A.J.: On the equivalence of control system and linearization of nonlinear systems. SIAM J. Control Optim. **11**, 670–676 (1973)
11. Cheng, D., Isidori, A., Respondek, W., Tarn, T.J.: Exact linearization of nonlinear with ouput. Math. Syst. Theory **21**, 63–68 (1998)
12. Kosoko, B.: Fuzzy systems as universal approximators. IEEE Trans. Comput. **43**(11), 1329–1333 (1989)
13. Amerongen, V.J.: Intelligent, Control (Part 1)-MRAS, Lecture Notes. University of Twente, The Netherlands, March (2004)
14. Abbas, H., Sajjad, S.Q.: Sliding mode control of coupled tank liquid level control system. In: IEEE 10th International Conference on Frontires of Information Technology (2012)

Improved LDPC Iterative Decoding Algorithm Based on the Reliable Extrinsic Information and Its Histogram

Nguyen Anh Tuan[1(✉)] and Pham Xuan Nghia[2]

[1] University of Information and Communication Technology, Thai Nguyen, Viet Nam
natuan@ictu.edu.vn
[2] Le Quy Don Technical University, Ha noi, Viet Nam

Abstract. In this paper we proposed a new method to prevent propagating errors due to passing the error Extrinsic information between nodes during the iterative decoding of low density parity check codes based on reliable Extrinsic information. Moreover, we also proposed a new method to analyze the convergence of the iterative LDPC decoding by using the histogram of Extrinsic information.

Keywords: LDPC decoding · Convergence of decoding · Reliable extrinsic information

1 Introduction

The convergence of iterative LDPC decoding processes are analyzed by the Density Evolution (DE) algorithm was proposed by Richardson et al. [1] or the Extrinsic Transfer Exit Chart devised by ten Brink [2]. Those above method help us in predicting the convergence of the LDPC codes and base on it we will decide the number of iterations used for decoding the LDPC codes. In this paper we introduce a novel method to predict the convergence based on analyzing the histogram of the Extrinsic information and also based on this histogram we will propose a new decoding method depended on the reliable Extrinsic information which are transferred between nodes during the LDPC decoding process.

2 A Novel Method to Predict the Convergence of the Iterative LDPC Decoding Process

The probabilistic LDPC decoding process is provided in [3] as following:

Based on the received soft values y_j at the output of the channel, the intrinsic probability of the jth bit being a binary 1 or binary 0 can be calculated as:

$$P(y|s_1) = \frac{1}{\sqrt{\pi N_0}} e^{\frac{-(y+\sqrt{E_b})^2}{N_0}}; P(y|s_0) = \frac{1}{\sqrt{\pi N_0}} e^{\frac{-(y-\sqrt{E_b})^2}{N_0}} \quad (1)$$

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_29

where, y and N_0 denotes the received soft channel output value and the power of channel noise, respectively.

The $P^1_{i,j}$ values the probability equals to 1of the neighbouring non-zero entries of the Equivalent Parity Check Matrix $\boldsymbol{H}_e$ are initialized by the $p(y|s_1)$ in Eq. (1).

The Extrinsic information $LR_{i,j}$ values corresponding to each non-zero entry in a given row of the $\boldsymbol{H}_e$ are updated as the below equation:

$$LR_{i,j} = \frac{1 + \prod_{l \in \{Ci\}, l \neq j} (1 - 2P^1_{i,j})}{1 - \prod_{l \in \{Ci\}, l \neq j} (1 - 2P^1_{i,j})} \tag{2}$$

where M, N are the number of rows and columns of the $\boldsymbol{H}_e$.

The probability ratio values corresponding to each non-zero entry in a given column of the $\boldsymbol{H}_e$ are updated:

$$PR_{i,j} = \frac{1 - P^1_j}{P^1_j} \prod_{k \in \{R_j\}, k \neq j} LR_{k,j} \tag{3}$$

The overall a posteriori probability ratio of the jth coded $PR(x_j)$ is updated as following:

$$PR(x_j) = \frac{1 - P^1_j}{P^1_j} \prod_{i \in \{R_j\}} LR_{i,j} with j = 1...N \tag{4}$$

The $P^1{}_{i,j}$ value corresponding to each non-zero entry of the $\boldsymbol{H}_e$ is updated according to $1/(1 + PR_{i,j})$, where $PR_{i,j}$ represents the updated values.

Based on the $PR(x_j)$ values updated in step 5, a tentative hard decision is made and this tentatively decoded codeword is multiplied with $\boldsymbol{H}^T$.

If the resultant syndrome vector is an all-zero vector, we declare a legitimate codeword has been found and the iterative decoding process is terminated.

By contrast, if the syndrome vector is not an all-zero vector and the maximum number of LDPC iterations is reached, we will declare a decoding failure and output the tentatively decoded codeword.

If the maximum affordable complexity has not been exhausted, go back to step 3.

Assuming that probabilities of the input bit having "1" and "0" values are equal each others. This means that $p(s_1) = p(s_0) = 1/2$. The error condition probability to receive transmitted s_1 and s_2 is given in the following equations:

$$P(e|s_1) = \frac{1}{\sqrt{\pi N_0}} \int_{-\infty}^{+\infty} e^{\frac{-(y-\sqrt{E_b})^2}{N_0}} dy = \frac{1}{2} erfc\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{5}$$

$$P(e|s_0) = \frac{1}{\sqrt{\pi N_0}} \int_{-\infty}^{+\infty} e^{\frac{-(y+\sqrt{E_b})^2}{N_0}} dy = \frac{1}{2} erfc\left(\sqrt{\frac{E_b}{N_0}}\right) \tag{6}$$

where: $erfc(x) = \frac{2}{\sqrt{\pi}} \int_{x}^{+\infty} e^{x^2} dx$ is the error compensating function. The error probability to fail to receive a transmitted bit is calculated as following:

$$P_b = \frac{1}{2} \mathrm{er} fc \left(\sqrt{\frac{E_b}{N_0}} \right) \tag{7}$$

From Eqs. (2), (5), (6) and (7) we can see that a single error received bit can be distributed to many other coded bit via the exchanging extrinsic information between nodes of the Tanner graph [4]. This distribution is very fast when the Eb/N0 is small and this error distribution causes the error avalanche. When the Eb/N0 value is high enough the error propagation is reduced, but this will delay the convergence of the LDPC decoding process and causes the error floor. To prevent this issue we will analyze the distribution of extrinsic information $LLR_{i,j}$ values (Log Likelihood Ratio) passed between nodes during the iterative LDPC decoding with the different the number of decoding iterations and Eb/N0 values in the next section.

To lead to the novel method predicting the convergence of the iterative LDPC decoding process we will analyze the distribution of $LLR_{i,j}$ values via the number of decoding iterations. We will simulate the histogram of $LLR_{i,j}$ values with the different parameters (The size of LDPC code word: 60, 120; code rate: $\frac{1}{2}$ Number of code words: 1000; Eb/N0 ratio: 4 dB; Modulation: BPSK). The LDPC is used in this simulation having the parity check matrix structure and using the decoding method proposed.

The distribution of the extrinsic information values after 2, 4 and 15 decoding iterations are plotted in the Figs. 1, 2 and 3. The transmission channel is the AWGN channel, the modulation is BPSK and the Eb/N0 = 4 dB. Observing the Figs. 1, 2 and 3, the distribution of the extrinsic information $LLR_{i,j}$ is changed via different numbers of decoding iterations. Those $LLR_{i,j}$ values are expanded toward two sides of the horizontal axes when the number of decoding iteration

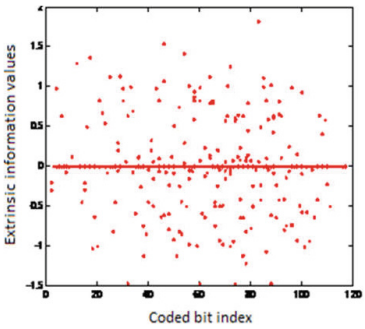


Fig. 1. The distribution of extrinsic information $LLR_{i,j}$ at Eb/N0 = 4 dB, Itermax = 2.

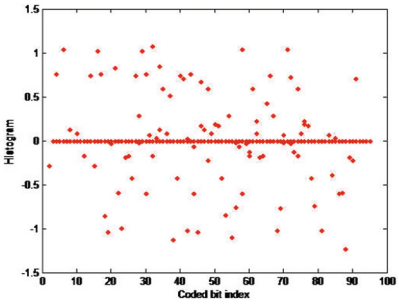


Fig. 2. The distribution of extrinsic information $LLR_{i,j}$ at Eb/N0 = 4 dB, Itermax = 4.

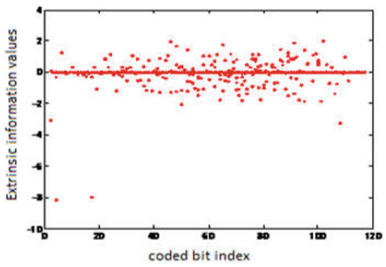


Fig. 3. The distribution of extrinsic information $LLR_{i,j}$ at Eb/N0 = 4 dB, Itermax = 15.

changes from 2 to 4, but most of them are converged around the horizontal axis at the 15th iteration. We can identify as following:

- At the number of decoding iterations equals to 2, most $LLR_{i,j}$ values concentrate near to the horizontal axis and when increasing the number of decoding iterations those values will be expanded to two sides of the horizontal axis as observed in the Fig. 2. However, when increasing the number of decoding iterations to 15 those above values will be converged back around the horizontal axis. This means that with the number of decoding iterations higher than 15 the values of the extrinsic information will be not so much increased. In the other word, there is no more valuable gain when increasing the number of decoding iterations over 15.
- There are quite a lot $LLR_{i,j}$ values equal to zero at different decoding iterations. This means that existing a lot of nodes not involved to the extrinsic information transferring process. This is caused because of the He having low density. The He having low density will prevent the error propagating during the LDPC decoding iteration. However, this also creates the error floor issue in decoding LDPC codes.

– By observing the distribution of extrinsic information values it is also provide for us a new method to analyze the convergent of the LDPC decoding having the same utility in comparison with the EXIT chart (Extrinsic Information Transfer) [5] or Density Evolution [6] methods. In our simulation, the $LLR_{i,j}$ values will be reduce toward the horizontal axis after the 15th iteration at Eb/N0 = 4 dB. This means that the LDPC decoding is almost converged after 15 decoding iterations. We will stop the decoding process after the 15th iteration at Eb/N0= 4 dB instead of continuing to iterate more the LDPC decoding. This help to reduce a lot the complexity of the decoding process.

By observing the distribution of the extrinsic information values $LLR_{i,j}$ at the different Eb/N0 ratios we also can improve the BER performance of LDPC codes by using the reliable $LLR_{i,j}$ values as presented in the following section.

3 A Method to Improving the Performance of LDPC Codes by Using the Reliable Extrinsic Information Values During the Iterative Decoding

Figures 4 and 5 are the distribution of the information values versus different Eb/N0 values at the number of decoding iterations equals to 2. As observing in the Figs. 4 and 5 we can notice that:

– At the low Eb/N0 values, the error transferring probability increases from the first to the second decoding iterations and then decreases from the second to the 15th iterations. Therefore, at the low Eb/N0, if we increase the number of decoding iterations to more than 2 times the BER will be increase, accordingly. The error increases because the iterative decoding process propagates errors via passing the error extrinsic information from one node to other related nodes.
– When the Eb/N0 values increase the extrinsic information values $LLR_{i,j}$ are also increase and their distribution will be expanded to two sides of the horizontal axis as seen in the above figures. At the high enough Eb/N0, the $LLR_{i,j}$ values are more reliable.

With the number of decoding iterations is bigger than 2 such as 15 times, the distribution of extrinsic information values $LLR_{i,j}$ are expanded to two side of zero axis. There are not so many abnormal values. Both the error and correct extrinsic information are propagated after each decoding iteration hence we could not identify the reliable extrinsic information values.

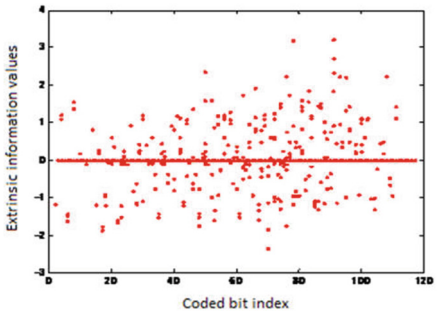


Fig. 4. The distribution of extrinsic information $LLR_{i,j}$ values at Eb/N0 = 2 dB after the 2nd iteration.

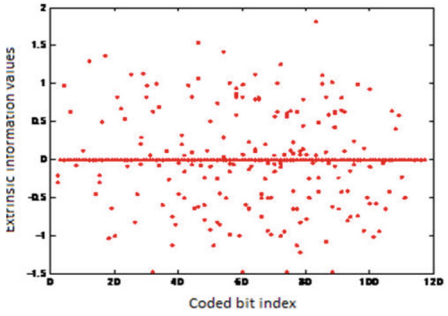


Fig. 5. The distribution of extrinsic information $LLR_{i,j}$ values at Eb/N0 = 4 dB after the 2nd iteration.

With the number of decoding iterations equals to 2, the $LLR_{i,j}$ values are distributed very close to the horizontal axis. Most of them are smaller than ±1.5. There are some values are bigger than ±1.5. We can say at the number of iterations equals to 2 the error extrinsic information is prevented from propagating to different nodes.

We need to consider to choice the right threshold at the as small as possible number of decoding iterations to prevent the error propagating issue in advance and also to reduce the total complexity of the iterative decoding. The decoding process in Fig. 6 is explained as following:

- At the first step: The soft bit yi from the demodulator are passed to the input of the decoder.
- The initial ${P^1}_j$ values are set to these soft bit values.
- Calculating the extrinsic information ratio $LLR_{i,j}$ values and check with the given threshold.
- If those values are satisfied the condition: $|LLR_{i,j}| \leq \pm 1.5$ reset these values to zero and updating the values $PR_{i,j}$ with the Eq. (3).
- If it is not, updating the values $PR_{i,j}$ with the Eq. (3).

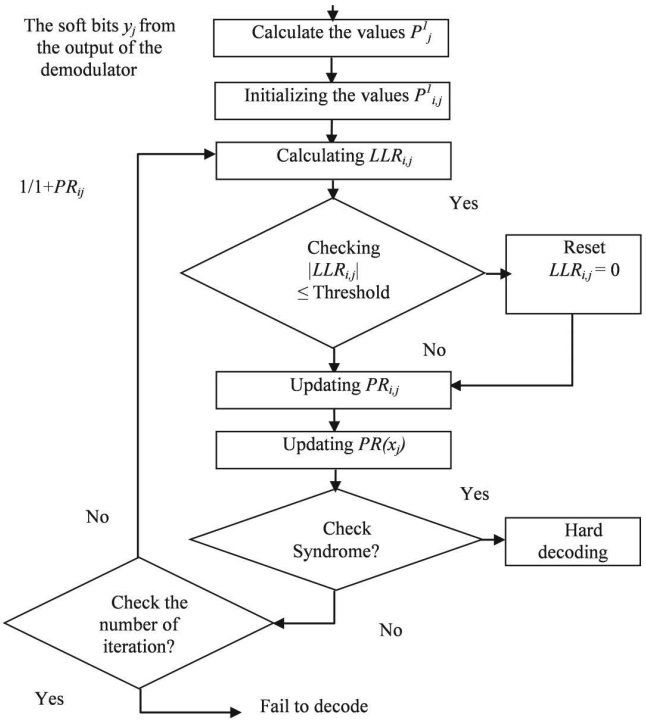


Fig. 6. The iterative decoding process based on reliable extrinsic information.

- Then updating the values $PR(x_j)$ with the Eq. (4).
- Checking the soft syndrome, if it is satisfied then pass the $PR(x_j)$ values to hard bit decoder and get the hard bits at the output. If it is not satisfied feeding back the value $1/(1 + PRi, j)$ to establishing the probabilities $P^1{}_{i,j}$ and continue with the next decoding processes. The simulation results of the novel decoding method are shown in the next section of this paper.

4 Simulation Result

The simulation parameters are listed in the Sect. 3. The LDPC is used in this simulation having the parity check matrix structure and using the BPA-EMS (Belief Propagation Algorithm based on Equivalent Parity Check Matrix and Minimum Weight of Syndrome) and BPA (Belief Propagation Algorithm) decoding method.

Figures 7 and 8 are the simulation BER performance versus Eb/N0 of LDPC codes using the BPA-EMS and BPA decoding method and our proposed method which uses the BPA-EMS and BPA decoding method based on the reliable extrinsic information to decode LDPCs after 10 decoding iterations.

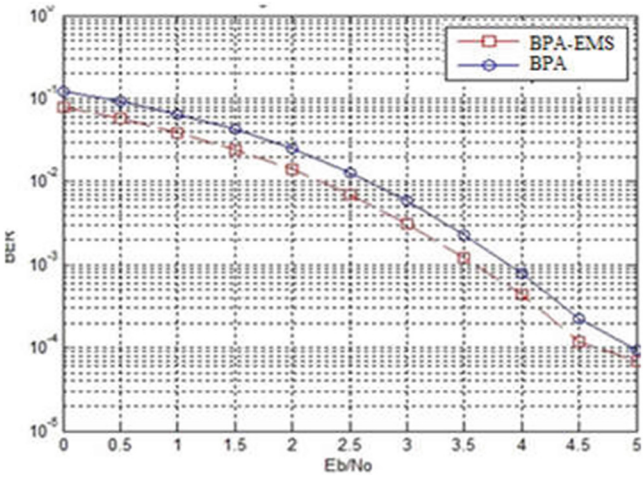


Fig. 7. The BER performance of LDPCs using the BPA-EMS and BPA decoding methods, 10 iterations, modulation BPSK in AWGN channel.

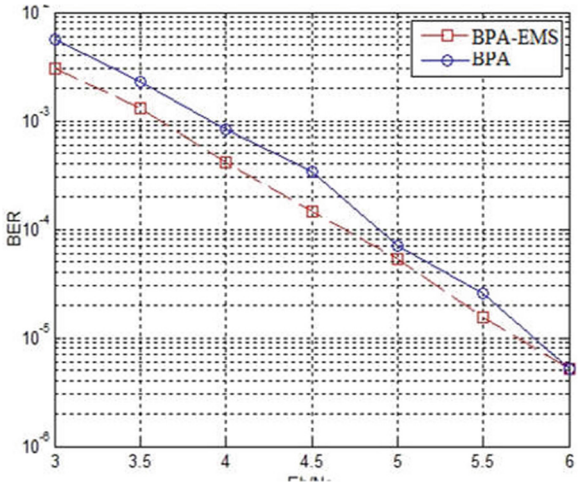


Fig. 8. The BER performance of LDPCs using the BPA-EMS and BPA decoding methods based on the reliable extrinsic information, number of iterations equals to 10 in AWGN channel.

As we can see in Figs. 7 and 8, LDPCs using the BPA-EMS and BPA decoding methods require the $Eb/N0 \geq 0.5\,\text{dB}$ to achieve the $BER = 10^{-4}$, while if LDPCs using our proposed decoding method require only 4.5 dB.

Figures 9 and 10 are the simulation BER performances versus $Eb/N0$ of LDPCs using the BPA-EMS and BPA decoding methods and our proposed after 15 decoding iterations. To archive the same $BER = 10^{-6}$ LDPCs using the BPA-

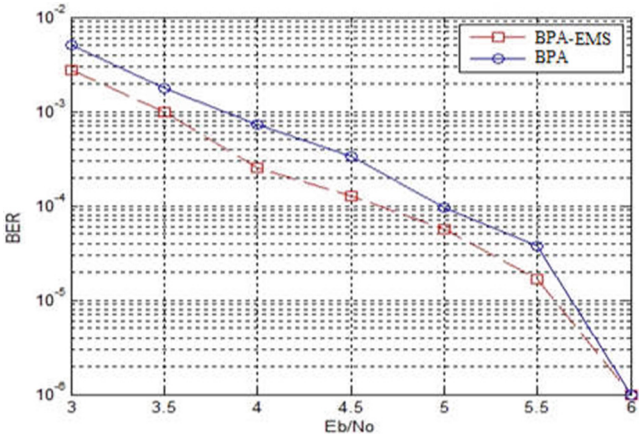


Fig. 9. The BER performance of LDPCs using the BPA-EMS and BPA decoding methods, the number decoding iterations is 15, modulation BPSK in AWGN channel.

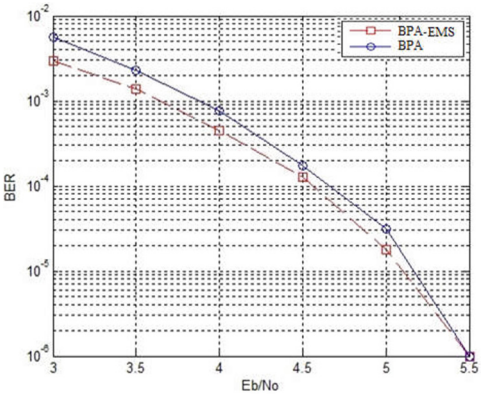


Fig. 10. The BER performance of LDPCs using the BPA-EMS and BPA decoding methods based on the reliable extrinsic information, the number of decoding iterations is 15.

EMS and BPA decoding methods require up to $Eb/N0 = 6\,\mathrm{dB}$, while LDPCs using our proposed method only need $Eb/N0 = 5.5\,\mathrm{dB}$.

5 Conclusion

In this paper we proposed our novel contributions which are a new method to analyze the convergence behavior of LDPC decoding process and an improved decoding method based on reliable extrinsic information to limit the error propagation during the iterative decoding of LDPCs. By using two methods proposed in this paper, the BER versus $Eb/N0$ performance of LDPCs gains 0.5 dB and

the complexity of the LDPC decoding process is also reduced a lot due to predicting the optimal number of decoding iterations. In the coming research we will concentrate to optimize these two methods to achieve better performances of LDPCs.

References

1. Chung, S.Y., Richardson, T.J., Urbanke, R.L.: Analysis of sum-product decoding of low densityparity check codes using a Gaussian approximation. IEEE Trans. Inf. Theory **47**, 599–618 (2001)
2. ten Brink, S.: Convergence behavior of iteratively decoded parallel concatenated codes. IEEE Trans. Commun. **49**, 1727–1737 (2001)
3. Hanzo, L., Liew, T.H., Yeap, B.L., Ng, S.X.: Turbo Coding, Turbo Equalisantion and Space Time Coding for Transmission Over Fading Channels. Wiley and IEEE (2002)
4. Tanner, R.M.: A recursive approach to low complexity codes. IEEE Trans. Inf. Theory **27**(5), 533–547 (1981)
5. Kollu, J., Hamid, J.: On the EXIT chart analysis of low density parity check codes. In: IEEE Global Telecommunications Conference, vol. 3 (2005)
6. Chen, J., Fossorier, M.P.C.: Density evolution for two improved BP-based decoding algorithms of LDPC codes. Commun. Lett. **6**, 208–210 (2002). IEEE

Improving Control Quality for Two-Wheeled Mobile Robot

Gia Thi Dinh(✉), Nguyen Duy Cuong, and Nguyen Dang Hao

Thai Nguyen University of Technology, 3/2 Street, Thai Nguyen City, Vietnam
{giadinh2206,nguyenduycuong,ndanghao}@tnut.edu.vn
http://www.tnut.edu.vn

Abstract. An approach for improving quality of motion control and stabilizing for two-wheeled mobile robot which is affected by internal and external distubances is presented in this paper. An output feedback controller using backstepping combined with high gain observer (HGOs) technique is designed. The decoupling technique for coupling system of two-wheeled mobile robot is also applied. Results of simulation show the effectiveness of the designed controller.

Keywords: Two-wheeled mobile robot · Underactuated nonlinear system · High gain observer

1 Introduction

Two-wheeled mobile robot with an inverted pendulum (TWMR), see Fig. 1, basically consists of a pair of identical actuators assembled on respective wheels, a chassis and an inverted pendulum. TWMR is balanced when the tilt angle ϕ of the pendulum is equal to zero or represents the inverted position. Conventionally the control objective is chosen such that the displacement x, heading angle ψ and the tilt angle ϕ asymptotically track the desired values, x_d, ψ_d and zero, respectively. It can be seen that there are only two inputs acting on the two wheels whereas there are three degrees of freedom to be controlled. Thus, TWMR can be classified as one kind of under-actuated systems [1,2].

Motion control for TWMR is one of the most popular problems used for testing nonlinear control algorithms. Various solutions for this issue ranging from linear to nonlinear controllers have been left out. Without concerning the disturbances and control of the heading angle, the authors in [3–5] use an energy approach to force the pendulum mass approaching the upper equilibrium point and swinging around it. The authors in [6,7] employ techniques based on Lyapunov approach to design a control law for keeping TWRM balancing at the equilibrium point where the authors in [8] apply a technique of nested saturations in order to stabilize the pendulum at upper equilibrium point. Continuing development and extension under the influence of external disturbances and control of the heading angle, the author [9] use an adaptive disturbance observer and the nested saturation combining with the backstepping technique to reject

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_30

the effect of external disturbances and keep the pendulum at vertical position. Intelligent controls [10,11], an on-line adaptive fuzzy approximation and multi-layer neural networks, are well-known techniques but in real system, it is difficult to choose satisfying fuzzy rule number. Moreover both techniques desire great computation. The authors in [12] develop an adaptive controller to eliminate the effect of disturbance but the quality of performance and robustness should continue to improve. This problem is solved by the author in [13]. However, when the frequency of disturbance is increased, the performance of system should be concerned. The above discussions motivate contributions of this paper.

2 Mathematical Model and Control Objective

This section presents the mathematical model of TWMR and proposes a control objective for designing control later.

2.1 Mathematical Model of TWMR

The physical model and diagram of forces and moments acting on TWMR are shown in Fig. 1. The parameters and variables of TWMR are described in the Table 1. Using the Lagrangian method, the equation of motion of TWMR is illustrated in (1) [13].

$$\begin{aligned}
\ddot{x} &= \frac{J_\phi}{\Omega_0 R}\left(T_L + T_R\right) - \frac{gm^2l^2}{\Omega_0}\sin(\phi)cos(\phi) \\
&\quad + \frac{J_\phi ml}{\Omega_0}\dot{\phi}^2\sin(\phi) + \frac{J_\phi}{\Omega_0}\left(f_{dR} + f_{dL}\right), \\
\ddot{\phi} &= -\frac{mlcos(\phi)}{\Omega_0 R}\left(T_L + T_R\right) - \frac{mlcos(\phi)}{\Omega_0 R}\left(f_{dR} + f_{dL}\right) \\
&\quad + \frac{M_x mgl}{\Omega_0}\sin(\phi) - \frac{m^2l^2}{\Omega_0}cos(\phi)\sin(\phi)\dot{\phi}, \\
\ddot{\psi} &= \frac{D}{2J_\psi R}\left(T_L - T_R\right) + \frac{D}{2J_\psi}\left(f_{dL} - f_{dR}\right).
\end{aligned} \tag{1}$$

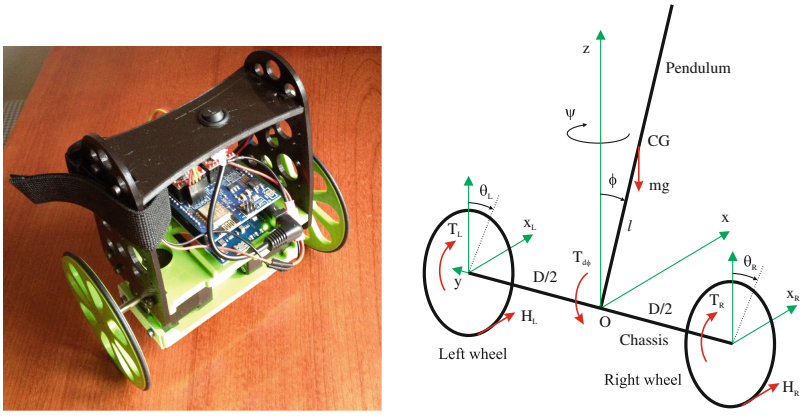


Fig. 1. Physical model and diagram of forces and torques acting on TWMR.

Table 1. TWMR parameters and variables

F_L, F_R	Interacting forces between the left and right wheels and the chassis
H_L, H_R	Friction forces acting on the left and right wheels
T_L, T_R	Torques provided by wheel actuators acting on the left and right wheels
f_{dL}, f_{dR}	External forces acting on the left and right wheels
θ_L, θ_R	Rotation angles of the left and right wheels
x_L, x_R	Displacements of the left and right wheels along the x-axis
ϕ	Tilt angle of the pendulum
ψ	Heading angle of the TWMR
x	Displacement of the TWMR along the x-axis
M_w	Mass of the wheel
J_w	Moment of inertia of the wheel with respect to the y-axis
R	Radius of the wheel
m	Mass of the pendulum
g	Gravity acceleration
l	Distance from the point O to the center of gravity, CG, of the pendulum
D	Distance between the left and right wheels along the y-axis
M	Mass of the chassis
J_c	Moment of inertia of the chassis about the y-axis
J_v	Moment of inertia of the chassis and pendulum about the z-axis
J_p	Moment of inertia of the pendulum about the z-axis

where

$$\begin{aligned} J_\phi &= \left(ml^2 + J_p + J_c\right) \\ M_x &= 2\left(M_\mathrm{w} + \frac{J_\mathrm{w}}{R^2}\right) + (M+m) \\ J_\psi &= \frac{D^2}{2}\left(M_\mathrm{w} + \frac{J_\mathrm{w}}{R^2}\right) + J_v \\ \Omega_0 &= M_x J_\phi - m^2 l^2 cos^2(\phi) \end{aligned} \tag{2}$$

Since $M_x J_\phi > m^2 l^2 cos^2(\phi)$, Ω_0 is positive for all $\phi \in \mathbb{R}$.

2.2 Control Objective

Assuming that at the initial time t_0, the inverted pendulum of TWMR is strictly in the upper half plane, i.e., there exists a strictly positive constant c_0 such that $|\phi(t_0)| \leq c_0$ and that the external forces f_{dL} and f_{dR} are unknown but bounded. The control objective is to design control laws, T_L and T_R, forcing TWMR tracking a series of desired way-points and stabilizing its inverted pendulum at the vertical position. In particular, the control laws are designed for T_L and T_R to guarantee:

$$\begin{aligned} &\lim_{t\to\infty}\left(x\left(t\right)-x_d\left(t\right)\right)=0,\\ &\lim_{t\to\infty}\left(\psi\left(t\right)-\psi_d\left(t\right)\right)=0,\\ &\lim_{t\to\infty}\phi\left(t\right)=0, \end{aligned} \tag{3}$$

for all $t \geq t_0 \geq 0$, $x_d(t)$ and $\psi_d(t)$ are reference trajectories of $x(t)$ and $\psi(t)$, respectively. For the TWMR tracking a series of desired way-points, $\ddot{x}_d(t) = 0$ is chosen.

3 Control Design

From the equation of motion of TWMR in (1), using the backstepping technique [14] combined with the high gain observer (HGOs) [15] designs control for TWMR satisfying the control objective in Sect. 2.2. First, a disturbance observer is designed, then, the dynamics of TWMR is separated into two subsystems, ψ-subsystem and (x, ϕ)-subsystem, which control by $T_L - T_R$ and $T_L + T_R$, respectively. The diagram of control design is shown in Fig. 2.

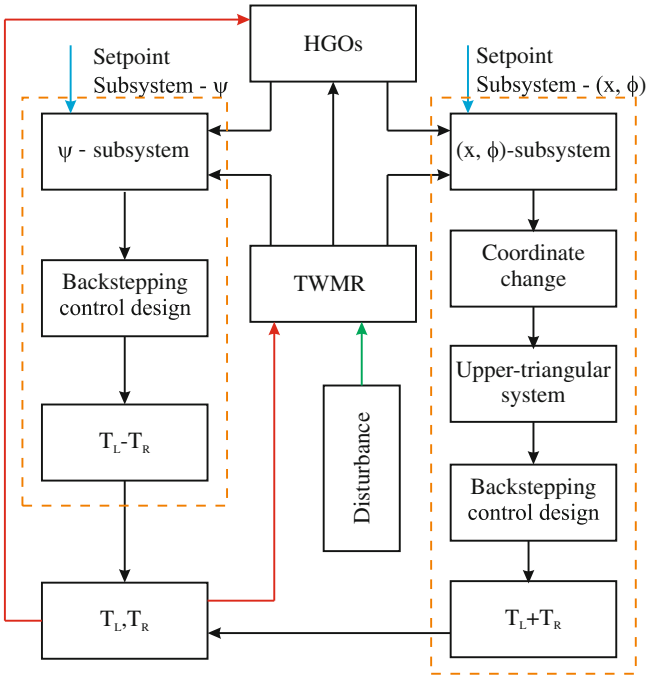


Fig. 2. Diagram of the control design.

3.1 Disturbance Observer

The dynamics of TWMR in (1) can be rewritten in the general form as in (4)

$$\begin{aligned} \dot{\boldsymbol{x}}_1 &= \boldsymbol{x}_2, \\ \dot{\boldsymbol{x}}_2 &= \boldsymbol{f}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{u}) + \boldsymbol{d}, \\ \boldsymbol{y} &= \boldsymbol{x}_1. \end{aligned} \tag{4}$$

Where $\boldsymbol{x}_1 = [x\, \phi\, \psi]^T$ and $\boldsymbol{x}_2 = [\dot{x}\, \dot{\phi}\, \dot{\psi}]^T$ are states. $\boldsymbol{u}$ is control input. $\boldsymbol{f}(\boldsymbol{x}_1, \boldsymbol{x}_2, \boldsymbol{u})$ is a function of $\boldsymbol{x}_1$, $\boldsymbol{x}_2$ and $\boldsymbol{u}$. $\boldsymbol{d} = [d_x\, d_\phi\, d_\psi]^T$ is unknown but bounded disturbance where $d \leq C_{d1}$ and $\dot{d} \leq C_{d2}$. Using HGOs technique, the disturbance observer is chosen as in (5).

$$\begin{aligned} \dot{\hat{\boldsymbol{x}}}_1 &= \hat{\boldsymbol{x}}_2 + h_1(\boldsymbol{y} - \hat{\boldsymbol{x}}_{\boldsymbol{1}}), \\ \dot{\hat{\boldsymbol{x}}}_2 &= \boldsymbol{f}(\boldsymbol{x}_1, \boldsymbol{u}) + h_2(\boldsymbol{y} - \hat{\boldsymbol{x}}_{\boldsymbol{1}}) + \hat{\boldsymbol{d}}, \\ \hat{\boldsymbol{d}} &= h_3(\boldsymbol{d} - \hat{\boldsymbol{d}}). \end{aligned} \tag{5}$$

Where h_1, h_2 and h_3 are positive constants.

The observer errors and their derivatives are shown in (6) and (7)

$$\begin{aligned} \tilde{\boldsymbol{x}}_1 &= \boldsymbol{y} - \hat{\boldsymbol{x}}_1, \\ \tilde{\boldsymbol{x}}_2 &= \boldsymbol{x}_2 - \hat{\boldsymbol{x}}_2, \\ \tilde{\boldsymbol{d}} &= \boldsymbol{d} - \hat{\boldsymbol{d}}. \end{aligned} \tag{6}$$

$$\begin{aligned} \dot{\tilde{\boldsymbol{x}}}_1 &= \hat{\boldsymbol{x}}_2 + h_1\tilde{\boldsymbol{x}}_{\boldsymbol{1}}, \\ \dot{\tilde{\boldsymbol{x}}}_2 &= -h_2\tilde{\boldsymbol{x}}_{\boldsymbol{1}} + \tilde{\boldsymbol{d}}, \\ \dot{\tilde{\boldsymbol{d}}} &= -h_3\tilde{\boldsymbol{d}} + \dot{\boldsymbol{d}}. \end{aligned} \tag{7}$$

The convergence of observer errors can be seen in the proof of Theorem 1.

3.2 Control Design

The backstepping technique [14] is employed to design control for two subsystems of TWMR.

ψ-Subsystem. The third equation of (1) with control input $T_L - T_R$ is rewritten in the state form as in (8)

$$\begin{aligned} \dot{\psi}_1 &= \psi_2, \\ \dot{\psi}_2 &= u_\psi + d_\psi. \end{aligned} \tag{8}$$

Where $\psi_1 = \psi$ and $\psi_2 = \dot{\psi}$ are states. $d_\psi = \frac{D(f_{dL} - f_{dR})}{2J_\psi}$ is unknown but bounded disturbance. $u_\psi = \frac{D(T_L - T_R)}{2RJ_\psi}$ is control input.

The errors are defined in (9)

$$\begin{aligned} \psi_{1e} &= \psi_1 - \psi_d, \\ \psi_{2e} &= \psi_2 - \alpha_\psi. \end{aligned} \tag{9}$$

Using the observer (5), the virtual control and control input are chosen as in (10)

$$\begin{aligned} \alpha_\psi &= -k_1\psi_{1e} + \dot{\psi}_d, \\ u_\psi &= -k_2\psi_{2e} - \psi_{1e} + \dot{\alpha}_\psi - \hat{d}_\psi. \end{aligned} \tag{10}$$

Where k_1 and k_2 are positive constants. α_ψ is a virtual control.

Substituting the virtual control and the control input in (10) into the time derivative of (9), the closed loop is shown in (11)

$$\begin{aligned} \dot{\psi}_{1e} &= -k_1\psi_{1e} + \psi_{2e} \\ \dot{\psi}_{2e} &= -k_2\psi_{2e} - \psi_{1e} + \tilde{d}_\psi \end{aligned} \tag{11}$$

The convergence of closed loop (11) sees in the proof of Theorem 1.

(x, ϕ)-Subsystem. The first and second equation of (1) are written in (12)

$$\begin{aligned} \ddot{x} &= u_{x\phi} + \tilde{d}_x \\ \ddot{\phi} &= -\tfrac{ml}{J_\phi} cos(\phi) u_{x\phi} + \tfrac{mgl\sin(\phi)}{J_\phi} + \tilde{d}_\phi \end{aligned} \tag{12}$$

where $d_x = \frac{J_\phi(f_{dR}+f_{dL})}{\Omega_0}, d_\phi = -\frac{mlcos(\phi)(f_{dR}+f_{dL})}{\Omega_0}, \tilde{d}_x = d_x - \hat{d}_x, \tilde{d}_\phi = d_\phi - \hat{d}_\phi, u_{x\phi} = \frac{1}{\Omega_0}\left[-gm^2l^2\sin(\phi)cos(\phi) + J_\phi\left(ml\dot{\phi}^2\sin(\phi) + \left(\frac{T_L}{R} + \frac{T_R}{R}\right)\right)\right] - \hat{d}_x, T_L + T_R = R\left(\frac{\Omega_0}{J_\phi}u_{x\phi} + \frac{\Omega_0}{J_\phi}\hat{d}_x + \frac{gm^2l^2\sin(\phi)cos(\phi)}{J_\phi} - ml\dot{\phi}^2\sin(\phi)\right)$.

To design control, the errors are defined

$$\begin{aligned} x_e &= x - x_d \\ \phi_e &= \phi \end{aligned} \tag{13}$$

The second derivative of (13) is illustrated in (14)

$$\begin{aligned} \ddot{x}_e &= u_{x\phi} + \tilde{d}_x \\ \ddot{\phi}_e &= a\sin(\phi_e) - b\cos(\phi_e)\, u_{x\phi} + \tilde{d}_\phi \end{aligned} \tag{14}$$

where $\ddot{x}_d = 0$, $a = \frac{mgl}{J_\phi}$ and $b = \frac{ml}{J_\phi}$. Using the coordinate change (15), the new form of (14) is shown in (16).

$$\begin{aligned} z_1 &= x_e + \tfrac{1}{b}\int_0^{\phi_e} \tfrac{1}{\cos(s)} ds, \\ z_2 &= \dot{x}_e - \tfrac{1}{b}\tfrac{\dot{\phi}_e}{\cos(\phi_e)}, \end{aligned} \tag{15}$$

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= \sin(\phi_{1e})\left(\tfrac{a}{b\cos(\phi_{1e})} + \tfrac{1}{b}\tfrac{\phi_{2e}^2}{\cos^2(\phi_{1e})}\right) \\ \dot{\phi}_{1e} &= \phi_{2e} \\ \dot{\phi}_{2e} &= v \end{aligned} \tag{16}$$

where $\phi_{1e} = \phi_e$, $\phi_{2e} = \dot{\phi}_e$, $v = a\sin(\phi_e) - b\cos(\phi_e)\, u_{x\phi}$. Applied backstepping technique, using z_2 to control z_1, ϕ_{1e} to control z_2, ϕ_{2e} to control ϕ_{1e} and v

to control ϕ_{2e}. Noting that $\lim_{\phi \to 0} \frac{\sin(\phi)}{\phi} = 1$ and $\frac{\sin \phi_{1e}}{\phi_{1e}} = \int_0^1 \cos(\phi_{1e} s)\, ds$, (16) is rewritten in (17)

$$\begin{aligned} \dot{z}_1 &= z_2 \\ \dot{z}_2 &= \phi_{1e} \Omega_\phi + \tilde{d}_z \\ \dot{\phi}_{1e} &= \phi_{2e} \\ \dot{\phi}_{2e} &= v + \tilde{d}_\phi \end{aligned} \tag{17}$$

where $\tilde{d}_z = \tilde{d}_x + \frac{\tilde{d}_\phi}{b \cos(\phi)}, \Omega_\phi = \left(\frac{a}{b \cos(\phi_{1e})} + \frac{1}{b} \frac{{\phi_{2e}}^2}{\cos^2(\phi_{1e})} \right) \int_0^1 \cos(\phi_{1e} s)\, ds.$

To design control, the errors are defined in (18)

$$\begin{aligned} z_{2e} &= z_2 - \alpha_{z_2} \\ \bar{\phi}_{1e} &= \phi_{1e} - \alpha_{\phi_1} \\ \bar{\phi}_{2e} &= \phi_{2e} - \alpha_{\phi_2} \end{aligned} \tag{18}$$

Choose the virtual control and control input as in (19)

$$\begin{aligned} \alpha_{z2} &= -c_1 z_1 \\ \alpha_{\phi 1} &= \frac{-c_2 z_{2e} + \dot{\alpha}_{z2} - z_1}{\Omega_\phi} \\ \alpha_{\phi 2} &= -c_3 \bar{\phi}_{1e} + \dot{\alpha}_{\phi 1} - z_{2e} \Omega_\phi \\ v &= -c_4 \bar{\phi}_{2e} + \dot{\alpha}_{\phi 2} - \bar{\phi}_{1e} \end{aligned} \tag{19}$$

Substituting the control (19) into the derivative of (18), the closed loop is obtained in (20)

$$\begin{aligned} \dot{z}_1 &= z_{2e} - c_1 z_1 \\ \dot{z}_{2e} &= \bar{\phi}_{1e} \Omega_\phi - c_2 z_{2e} - z_1 - \tilde{d}_z \\ \dot{\bar{\phi}}_{1e} &= \bar{\phi}_{2e} - c_3 \bar{\phi}_{1e} - z_{2e} \Omega_\phi \\ \dot{\bar{\phi}}_{2e} &= -c_4 \bar{\phi}_{2e} - \bar{\phi}_{1e} - \tilde{d}_\phi \end{aligned} \tag{20}$$

The control design has completed. The results are summarized in the Theorem 1.

Theorem 1. *With the control objective proposed in Sect. 2.2, using the disturbance observer (5) and applying the control inputs (10) and (19), the closed loop system (11) and (20) are practical stable around the origin. If $C_{d2} = 0$, they are asymptotically stable at the origin.*

Proof. It is easy to see that if the effect of $\tilde{\boldsymbol{d}}$ is not concerned, two first equations of (7) satisfy Hurwitz stable. With the third equation of (7), considering the time derivative of Lyapunov candidate $V_d = \frac{1}{2} \tilde{\boldsymbol{d}}^2$

$$\begin{aligned} \dot{V}_d &= -h_3 \tilde{\boldsymbol{d}}^2 + \tilde{\boldsymbol{d}} \dot{\boldsymbol{d}} \leq -\left(h_3 - \delta\right) \tilde{\boldsymbol{d}}^2 + \frac{1}{4\delta} \dot{\boldsymbol{d}}^2 \\ &\leq -2\left(h_3 - \delta\right) V_d + \frac{1}{4\delta} C_{d2}^2 \end{aligned} \tag{21}$$

where δ is a positive constant such that $(h_3 - \delta) > 0$. It can be seen that V_d converges to a sphere, center at the origin, radius $R_{V_d} = \frac{C_{d2}^2}{8\delta(h_3-\delta)}$ when $\boldsymbol{x}_1$ and $\boldsymbol{x}_2$ exist. Thus, $\tilde{\boldsymbol{d}}$ stabilizes to a sphere, center at origin, radius

$$R_{\tilde{d}} = \sqrt{\frac{1}{4\delta\left(h_3 - \delta\right)}} C_{d2}, \tag{22}$$

To prove the convergence of the closed loop (11) and (20), consider the following candidate

$$V = \frac{1}{2}\left(\psi_{1e}^2 + \psi_{2e}^2\right) + \frac{1}{2}\left(z_1^2 + z_{2e}^2 + \bar{\phi}_{1e}^2 + \bar{\phi}_{2e}^2\right) \tag{23}$$

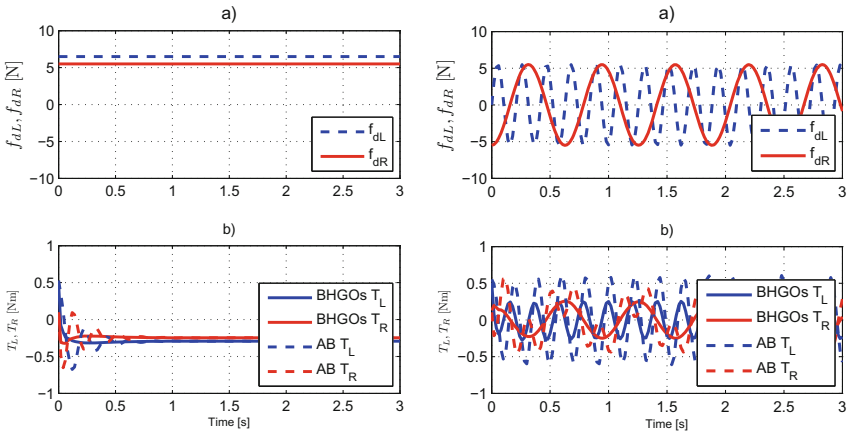
Taking the time derivative of (23) yields

$$\dot{V} = -k_1\psi_{1e}^2 - k_2\psi_{2e}^2 - c_1 z_1^2 - c_2 z_{2e}^2 - c_3\bar{\phi}_{1e}^2 - c_4\bar{\phi}_{2e}^2 \leq -kV \tag{24}$$

From the inequality (24), it can be concluded that the closed loop (11) and (20) are asymptotically stable at the origin. This implies that ψ_{1e}, ψ_{2e}, z_1, z_{2e}, $\bar{\phi}_{1e}$, and $\bar{\phi}_{2e}$ are bounded and $\lim_{t\to\infty}\left(\psi_{1e}, \psi_{2e}, z_1, z_{2e}, \bar{\phi}_{1e}, \bar{\phi}_{2e}\right) = 0$. The control inputs (10) and (19) are also bounded. These lead the boundedness of ψ_{1e}, ψ_{2e}, x_{1e}, x_{2e} , ϕ_{1e} and ϕ_{2e} and the errors satisfy $\lim_{t\to\infty}\left(\psi_{1e}, \psi_{2e}, x_{1e}, x_{2e}, \phi_{1e}, \phi_{2e}\right) = 0$ or the control objective proposed in (3) is achieved. The Theorem 1 is proved.

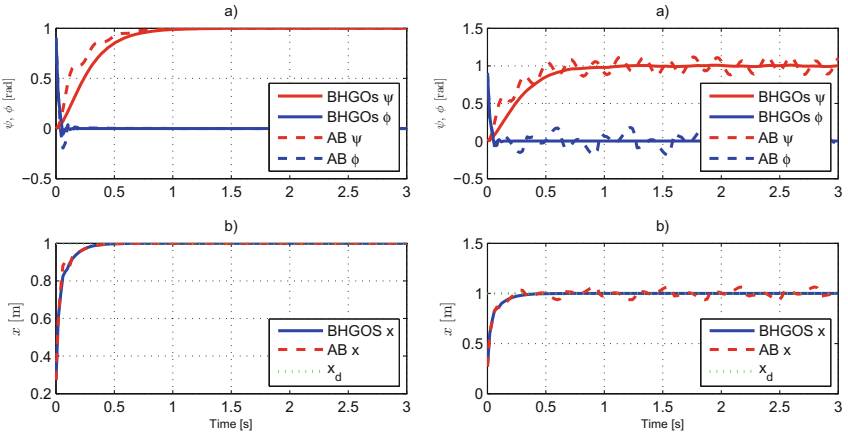
4 Simulation

To illustrate the effectiveness of the proposed controllers stated in Theorem 1. In the simulation, the parameters of TWMR are taken as $M = 0.35[kg]$,



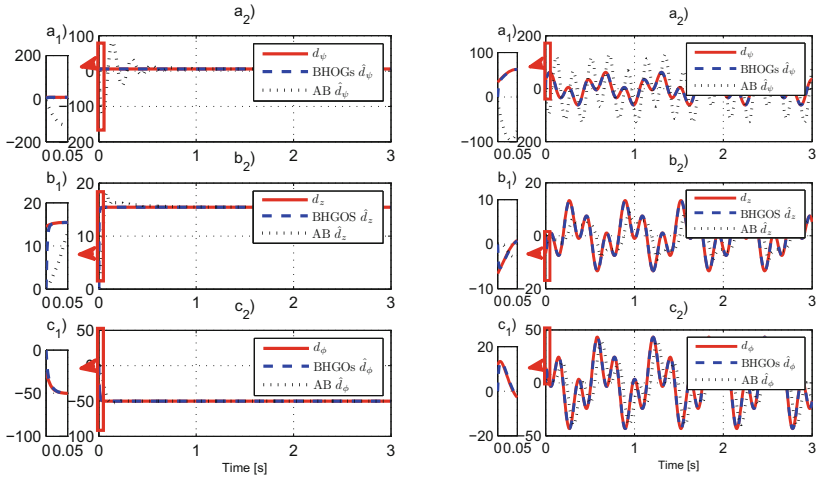
(a) Case 1. (b) Case 2.

Fig. 3. Disturbance and control input.



(a) Case 1. (b) Case 2.

Fig. 4. Performance results.



(a) Case 1. (b) Case 2.

Fig. 5. Disturbance and its estimation.

$g = 9.8[m/s^2]$, $M_w = 0.05[kg]$, $R = 0.045[m]$, $D = 0.18[m]$, $J_w = 0.000253$ $[kg.m^2]$, and $J_c = 0.00017[kg.m^2]$.

The initial conditions are taken as $x(0) = 0[m]$, $\dot{x}(0) = 0[m/s]$, $\phi(0) = 1[rad]$, and $\dot{\phi}(0) = 0.2[rad/s]$, $\psi(0) = 1[rad]$. The gain controls are chosen as $k_1 = 5$, $k_2 = 10$, $c_1 = 10$, $c_2 = 5$, $c_3 = 50$, and $c_4 = 40$. The reference values are taken as $x_d = 1[m]$ and $\psi_d = \pi/4[rad]$. The observer and update law gains are take as $h_1 = 100$, $h_2 = 200$, $h_3 = 300$ and $\gamma = 3000$.

To demonstrate the improvement of designed controller (BHGOs), the simulation is executed in two cases and in comparing with adaptive backstepping controller (AB). In the first case, the disturbance is chosen with constant value whereas in the second case, the disturbance is taken with sinusoidal oscillation of different frequency.

Case 1: $f_{dL} = 6.5$ [N], $f_{dR} = 5.5$ [N]. The disturbance and control input are shown in Fig. 3a. The response and observer of ψ, ϕ and x are plotted in Figs. 4a and 5a. It can be seen from Fig. 4a that the response of ψ, ϕ and x of AB and BHGOs have the same quality. However, the observer of AB has higher value than the observer of BHGOs. This has an effect on the transient performance. Figure 3a shows that the control input of AB has higher value and frequency comparing with BHGOs.

Case 2: $f_{dL} = 5.5\sin(30t)$ [N], $f_{dR} = 5.5\cos(10t)$ [N]. Using different frequency of disturbance, $30/\pi$ and $10/\pi$, The disturbance and control input are shown in Fig. 3b. The response and observer of ψ, ϕ and x are plotted in Figs. 4b and 5b. From Fig. 4b, it is clear to see that the performance of AB has some oscillations where the performance of BHGOs is still unchanged. This is because of that the observer of BHGOs is more accurate than AB. The stabilization of ϕ under the effect of disturbance improves the quality of control for TWMR.

5 Conclusion

A control design technique combined with the high gain observer to construct control for TWMR has been presented. The simulation results show that under the effect of disturbance, the performance of BHGOs has better quality than the accomplishment of AB, specially, when the frequency of disturbance is increased. Continuing research and applying the designed controller BHGOs in experimental work are the future work.

References

1. Do, K.D., Pan, J.: Control of Ships and Underwater Vehicles. Springer, London (2009)
2. Olfati-Saber, R., Megretski, A.: Controller design for a class of underactuated nonlinear systems. In: Proceedings of the 37th IEEE Conference on Decision and Control, vol. 4, pp. 4182–4187, December 1998
3. Astrom, K., Furuta, K.: Swinging up a pendulum by energy control. Automatica **36**(2), 287–295 (2000)
4. Chung, C.C., Hauser, J.: Nonlinear control of a swinging pendulum. Automatica **31**(6), 851–862 (1995)
5. Lozano, R., Fantoni, I., Block, D.J.: Stabilization of the inverted pendulum around its homoclinic orbit. Syst. Control Lett. **40**(3), 197–204 (2000)
6. Wei, Q., Dayawansa, W., Levine, W.: Nonlinear controller for an inverted pendulum having restricted travel. Automatica **31**(6), 841–850 (1995)

7. Mazenc, F., Praly, L.: Adding an integration and global asymptotic stabilization of feedforward systems. In: Proceedings of the 33rd IEEE Conference on Decision and Control, vol. 1, pp. 121–126, December 1994
8. Teel, A.: A nonlinear small gain theorem for the analysis of control systems with saturation. IEEE Trans. Autom. Control **41**, 1256–1270 (1996)
9. Do, K.D., Seet, G.: Motion control of a two-wheeled mobile vehicle with an inverted pendulum. J. Intell. Robot. Syst. **60**, 577–605 (2010)
10. Wu, Q., Sepehri, N., He, S.: Neural inverse modeling and control of a baseexcited inverted pendulum. Eng. Appl. Artif. Intell. **15**(34), 261–272 (2002)
11. Vapnik, V.: An overview of statistical learning theory. IEEE Trans. Neural Netw. **10**, 988–999 (1999)
12. Son, N.N., Anh, H.P.H.: Adaptive backstepping self-balancing control of a two-wheel electric scooter. Int. J. Adv. Robot. Syst., October 2014
13. Dinh, G.T., Cuong, N.D.: Robust adaptive control for two-wheeled mobile robot. In: VCCA-2015 (2015)
14. Krstic, M., Kanellakopoulos, L., Kokotovic, P.: Nonlinear and Adaptive Control Design. John Wiley and Sons, New York (1995)
15. Khalil, H.K., Praly, L.: High-gain observers in nonlinear feedback control. Int. J. Robust Nonlinear Control **24**(6), 993–1015 (2014)

Incorporation of Experience and Reference-Based Topic Trust with Interests in Social Network

Phuong Thanh Pham[1], Manh Hung Nguyen[2,3], and Dinh Que Tran[2(✉)]

[1] Thang Long University, Hanoi, Vietnam
ppthanh216@gmail.com
[2] Post and Telecommunication Institute of Technology, Hanoi, Vietnam
nmh.nguyenmanhhung@gmail.com, tdque@yahoo.com
[3] UMI UMMISCO 209 (IRD/UPMC), Hanoi, Vietnam

Abstract. Computational trust in the social media depends both on the interaction types and on the expertise, which exposes user's interests by tags on items such as books, articles, images etc. However, when there is not any interaction among users, such a computation is impossible. In this paper, we first propose a novel model of computational trust among users in social network to incorporate the interaction among users and semantics of topics based on tags posted by users. Both types of direct and indirect interaction via intermediate ones are utilized for computation in the model. Then, we introduce algorithms for computing trust among users in social network.

Keywords: Computational trust · Tag · Interest · Topic · Social network

1 Introduction

Virtual networks such as forums, blogs, wiki, social networks, etc. have facilitated users to interact with each other. By means of interaction, they may share information, opinions, exchange ideas or make friendship etc. In social networks, users utilize their own tags to annotate and organize items for searching or sharing. They also make use of tags for showing their viewpoint or opinions. Such tags are a kind of meta-data composed of keywords or terms to introduce bookmarks, article titles, comments of digital image etc. These tags have contributed to discovering knowledge for various real world applications such as recommender systems, searching engine, predicting customer opinions [1–3,8,10,11].

In the process of interaction, trust has become crucial for belief among partners to share or post their messages. The issue has attracted increasing research interests in computer science. There are various models of computational trust proposed in literature [4–7,9]. However, such computation approaches are mainly based on interaction among partners rather than the semantics of messages communicated in the network.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_31

The purpose of this paper is first to propose a computational model of trust by incorporating an interaction-based trust and interest-based expert level of users in social media. The former is based on temporal interaction given by Nguyen et al. [5] and the latter is constructed from user's interests described via tags. Then, we describe an algorithm for incorporation computation of trust among users in social network.

The remainder of this paper is organized as follows. Section 2 describes preliminaries for presenting the following section. Section 3 presents a computational model of incorporation of trust and user's expert level based on his interests. Section 4 is a conclusion.

2 Preliminaries

This section presents some concepts and notations, which are basis for the presentation in the next section. The computation of trust based on interaction among users will be a basis for constructing the concept of *topic trust*, which describes the expertise of users when they post tags of items.

- Each user in social media may be considered as an autonomous entity in the system. Let $\mathcal{U} = \{u_1, \ldots, u_m\}$ be a set of users.
- Let I_{ij} be a set of all interactions between u_i and u_j and $\mid I_{ij} \mid$ be the number of interactions. Interaction between users u_i and u_j is a set of transactions at an instant time, which occurs when u_i sends to u_j a message such as post, comment, opinions etc. In such an interaction, u_i might give a trust value on u_j.
- Tag is a brief piece of information dispatched from some user u_i to make a description or post information/idea/opinions on an item such as a paper, a book, a film, a thing and so on.
- Intuitively, when an user is interested in some topic t, he is willing to dispatch a tag about items on it. From such tags, we can construct a classification of all tags into the set of topics. We may use any technique for such a classification (Refer to [2] for more detail). Suppose that $\mathcal{T} = \{t_1, \cdots, t_n\}$ is a set of such topics. We denote $classifier(tags, Topic)$ the function for classifying tags into classes.

Denote $w = (w_1, w_2, ...w_{|I_{ij}|})$ to be the weight vector, in which $\mid I_{ij} \mid$ is the number of interaction among users u_i and u_j, such that $\sum_{k=1}^{|I_{ij}|} w_k = 1$ and $w_{k_1} \leq w_{k_2}$ if $k_1 \leq k_2$. This vector is used to describe the weighted evaluation of temporal trust. The constraints $w_{k_1} \leq w_{k_2}$ if $k_1 \leq k_2$ illustrates the fact that the recent evaluations of an user u_i on an user u_j is more important than the previous ones. The construction of the weight vector may be based on Regular Decreasing Monotone (RDM) function. Computing experience trust based on purely direct interaction proposed by Nguyen et al. is given in Algorithm 1, in which $generateW(k)$ is a function generating the weight vector w of size k (Refer to [5] for more detail).

Definition 1. *Experience trust of user u_i on user u_j, denoted $trust^{exp}(i,j)$, is defined by the formula*

$$trust^{exp}(i,j) = e_{ij} = t_{ij} * w = \sum_{k=1}^{|I_{ij}|} t_{ij}^k * w_k \quad (1)$$

where t_{ij} is the vector of temporal interaction trust of user u_i on its partner u_j and w is the vector of weighted interaction.

Algorithm 1. Experience trust of u_i on u_j via interaction

Input: The set of interactions of users u_i and u_j
Output: the trust of u_i on u_j, $calculateTrust_{exp}(i,j)$.

1: **for all** user u_i in the system **do**
2: **if** there is a new transaction k with a partner u_j **then**
3: $t_{ij}^k \leftarrow 0$ or 0.5 or 1 // 0: negative; 0.5: neutral; 1: positive
4: $t_{ij} \leftarrow t_{ij} \cup t_{ij}^k$ // add t_{ij}^k into t_{ij}
5: $t_{ij} \leftarrow sort(t_{ij})$ //re-sort t_{ij} on descending of time
6: $w \leftarrow generateW(k)$ //generate the weight vector w of size k
7: $e_{ij} \leftarrow \sum_{h=1}^{k} t_{ij}^h * w_h$ //update the experience trust
return $trust_{exp}(i,j)$

Definition 2. *A topic trust of u_i on u_j of t is a function $trust_{topic} : \mathcal{U} \times \mathcal{U} \times \mathcal{T} \rightarrow [0,1]$, in which $[0,1]$ is an unit interval of the real numbers. The value $trust_{topic}(i,j,t) = \alpha$ means that u_i (truster) trusts u_j (trustee) of topic t with respect to the degree α.*

3 Topic Trust based on Experience and Reference

3.1 Problem Statement

Given two users u_i and u_j, we would like to compute a topic trust of u_i on u_j of topic t. There are two cases. First, when there is a direct interaction between u_i and u_j, then the result trust is computed based on *experience trust*. Second, when there is no interaction between them, then trust is defined based on *neighbor*. It means that the computation is based on *reference topic trust* $trust_{topic}(k,j,t)$ for all u_k, who have direct interaction with both u_i and u_j. This section is devoted to presenting the steps of such a computation.

3.2 Topic Trust Based on Experience

This section presents definitions of concepts and algorithms for computing trust in the case there is interaction between users u_i and u_j.

In order to estimate the level of interest of an user u_i in a given topic t, we assume that:

- For each user $u_i \in \mathcal{U}$ and for all topics in $\mathcal{T}$, the bigger the number of tags in a topic $t \in \mathcal{T}$, the higher the level of interest of user u_i in the topic t, regarding other topics in $\mathcal{T}$.
- For each topic $t \in \mathcal{T}$ and for all users in $\mathcal{U}$, the bigger the number of tags that an user $u_i \in \mathcal{U}$ has dispatched in the topic t, the higher the level of interest of user u_i in the topic t, regarding other users in $\mathcal{U}$.

So we define the level of interest of u_i on topic t as follow.

Definition 3. *Suppose that n_t^i is the number of tags an user u_i has dispatched in some topic t. Then the level of interest of u_i on topic t is defined by the following formula*

$$interest_{topic}(i,t) = \frac{1}{2}\left(\frac{n_t^i}{\sum_{k\in\mathcal{T}} n_k^i} + \frac{n_t^i}{\sum_{u_k\in\mathcal{U}} n_t^k}\right) \tag{2}$$

Definition 4. *Suppose $interest_{topic}(i,t)$ is the level of interest of u_i on topic t. Then the level of expert of user u_i on the topic t is defined as follows*

$$expert(i,t) = interest_{topic}(i,t) \tag{3}$$

Based on the trust among users via experience and expert level of an user on some topic t, we can construct a trust of u_i on u_j of topic t based on topic-trust function as follows.

Definition 5. *A function $f : [0,1] \times [0,1] \to [0,1]$ is a topic-trust function iff it is a increasing monotonic w.r.t. each variable. Denote $f(x,y) = x \odot y$ and $\odot$ is called the topic-trust operator.*

It is easy to prove the following proposition.

Proposition 1. *The following functions $f : [0,1] \times [0,1] \to [0,1]$ are topic-trust ones:*

(i) $f(x,y) = x \times y$
(ii) $f(x,y) = max(x,y)$
(iii) $f(x,y) = min(x,y)$
(iv) $f(x,y) = w_1 \times x + w_2 \times y$, *where* w_1 *and* w_2 *are weights, i.e.,* $w_1 + w_2 = 1$

Definition 6. *Suppose that* $trust_{exp}(i,j) = \alpha$ *and* $expert(j,t) = \beta$*, then the trust of* u_i *on* u_j *of topic* t *is defined by the following formula:*

$$trust^{exp}_{topic}(i,j,t) = \alpha \odot \beta \tag{4}$$

where $\alpha \odot \beta$ *is a topic-trust operator.*

It is easy to prove the following propositions.

Proposition 2. *If* $expert(i,t) \leq expert(j,t)$*, then for all* u_k*,* $trust^{exp}_{topic}(k,i,t) \leq trust^{exp}_{topic}(k,j,t)$.

Proposition 3. *If* $trust_{exp}(i,j) \leq trust_{exp}(i,k)$*, then for all* t*,* $trust^{exp}_{topic}(i,j,t) \leq trust^{exp}_{topic}(i,k,t)$.

Proposition 2 confirms the intuition that the higher expert level an user has, the more trust he is. Proposition 3 indicates that the more interaction an user u_i has with another u_j, the more trust u_i on u_j is. Steps for computing topic trust with direct interaction among users u_i and u_j in the social network is given in Algorithm 2.

Algorithm 2. Experience Trust of u_i on u_j of topic t

Input: The set of topics $\mathcal{T} = \{t_1, t_2, ..., t_n\}$ and the set of users $\mathcal{U} = \{u_1, u_2, ..., u_m\}$ with tags

Output: the trust of u_i on u_j of topic t, $calculateTrust^{exp}_{topic}(i,j,t)$.

1: **for all** t in $\mathcal{T}$ **do**
2: **for all** $u_i \in \mathcal{U}$ **do**
3: $n^t_i \leftarrow numberOfTags(i,t)$ //Number of tags u_i post on topic t
4: $$expert(i,t) \leftarrow \frac{1}{2}\left(\frac{n^i_t}{\sum_{k\in\mathcal{T}} n^i_k} + \frac{n^i_t}{\sum_{u_k\in\mathcal{U}} n^k_t}\right)$$
5: $trust^{exp}_{topic}(i,j,t) \leftarrow trust^{exp}(i,j) \odot expert(j,t)$ **return** $trust^{exp}_{topic}(i,j,t)$

3.3 Topic Trust Based on Reference

This section is to describe the computation of topic trust of u_i on u_j when there is no direct interaction between them. If u_j is an user with which u_i has direct interaction, then u_j is called a *neighbor* of u_i. Then based on Algorithm 2, we can compute trust $trust^{exp}_{topic}(i,j,t)$ for all topics t. Since there is no any interaction between u_i and u_j, it is impossible to compute $trust^{exp}_{topic}(i,j,t)$.

The computation of trust of u_i on u_j may be determined by means of neighbors at various levels. If u_i has a direct interaction with u_j, then u_j is called a neighbor level 1 of u_i, denote *1-neighbor* of u_i. The concept of k-neighbor of u_i is defined as follows.

Definition 7. *An user u_j is a neighbor level k ($k \geq 2$) of user u_i, called k-neighbor of u_i, iff*

- *u_j has no direct interaction with any user which is a z-neighbor of u_i, for all $z \leq k-2$, and*
- *There is at least an user which is a (k-1)-neighbor of u_i has direct interaction with u_j.*

Denote $neighbor_k(i,j)$.

Definition 8. *Suppose that F_i^1 is a set of 1-neighbor of u_i. For all $k \geq 2$, F_i^k is defined as follows*

$$F_i^k = \{u_j \mid \text{there is at least one } u_v \in F_{k-1}^i \text{ interacting with } u_i \text{ and, there is not any user } u'_v \in F_z^i, (\forall z \leq k-2) \text{ interacting with } u_j\} \quad (5)$$

Definition 9. *Given users u_i and u_j. Denote $F_i^{1,j}$ is a set of 1-neighbors of u_i that has some interaction with u_j. For all $k \geq 2$, $F_i^{k,j}$ is defined as follows*

$$F_i^{k,j} = \{u_v \mid \text{there is at least an user } u_l \in F_{k-1,j}^i \text{ interacting with } u_i \text{ and, there is not any user } u'_l \in F_{z,j}^i, (\forall z \leq k-2) \text{ interacting with } u_j\} \quad (6)$$

Thus, we have a taxonomy of neighbors of u_i. F_i^1 are closest ones and satisfies conditions $neighbor_1(i,j)$ for all $u_j \in F_i^1$. And $F_i^{1,j}$ is a set of users u_k has some interaction with both a_i and a_j. Based on Definition 6, we can compute topic trust of u_i on u_k of topic t for all $u_k \in F_i^{1,j}$. And in turn, we can compute trust u_i on u_j via all u_k in $F_i^{1,j}$, which is called *reference-based topic trust*. The definition of topic trust based on reference is given in Definition 10.

Definition 10. *Let $F_{1,j}^i$ be a set of users that have interaction with both u_i and u_j. Suppose that $F_{1,j}^i$ is not a empty set. A reference-based topic trust of u_i on u_j of topic t is determined by the formula*

$$trust_{topic}^{ref}(i,j,t) = \frac{1}{\mid F_i^{1,j} \mid} \sum_{u_k \in F_i^{1,j}} trust_{topic}^{exp}(i,k,t) \odot trust_{topic}^{exp}(k,j,t) \quad (7)$$

where $\odot$ is a topic trust operator.

The computation of reference-based topic trust of u_i on u_j by means of his neighbors is described in Algorithm 3.

Algorithm 3. Reference Trust of u_i on u_j of topic t

Input: The set of topics $\mathcal{T} = \{t_1, t_2, ..., t_n\}$ and the set of users $\mathcal{U} = \{u_1, u_2, ..., u_m\}$
Output: the trust of u_i on u_j of topic t, $calculateTrust^{ref}_{topic}(i, j, t)$.

$F_i^{1,j} \leftarrow constructInteractionSet(i, j, \mathcal{U})$ //the set of all users interaction with both u_i and u_j
if $F_i^{1,j} \neq \emptyset$ **then**
  for all t in $\mathcal{T}$ **do**
    for all $u_k \in F_i^{1,j}$ **do**
      $t_i^k \leftarrow t_k^i + calculateTrust^{exp}_{topic}(i, k, t)$ //Algorithm 2
      $t_k^j \leftarrow t_j^k + calculateTrust^{exp}_{topic}(k, j, t)$ //Algorithm 2
      $t_i^j \leftarrow t_i^j + t_i^k \odot t_k^j$
  $trust^{ref}_{topic}(i, j, t) \leftarrow \frac{t_j^i}{|F_i^{1,j}|}$
return $trust^{ref}_{topic}(i, j, t)$
else return 0

4 Conclusions

In this paper, we have introduced a model of incorporation trust, which integrates the computational trust based on experience of direct interaction and expert level with semantics based on tags in social network. And in turn, the level of expert in this paper is based on interests of topics which is classified from tags. There are some unsolved open questions in our work. The first one is if the computation of incorporation trust depends on various techniques of classification. Second, whether the reference-based topic trust should be computed by means of 1-neighbors of a user himself or deeper ones. Third, if there is a critical difference when selecting the various topic-trust functions. The issues need to be investigated furthermore. We are currently performing experimental evaluation and comparing with other models on computing trust in social network. The research results will be presented in our future work.

References

1. Feng, W., Wang, J.: Incorporating heterogeneous information for personalized tag recommendation in social tagging systems. In: Proceedings of the 18th ACM SIGKDD International Conference on Knowledge Discovery and Data Mining, KDD 2012, pp. 1276–1284. ACM, New York (2012)
2. Gattani, A., Lamba, D.S., Garera, N., Tiwari, M., Chai, X., Das, S., Subramaniam, S., Rajaraman, A., Harinarayan, V., Doan, A.H.: Entity extraction, linking, classification, and tagging for social media: a wikipedia-based approach. Proc. VLDB Endow. **6**(11), 1126–1137 (2013)
3. Li, X., Guo, L., Zhao, Y.E.: Tag-based social interest discovery. In: Proceedings of the 17th International Conference on World Wide Web, WWW 2008, pp. 675–684. ACM, New York (2008)
4. Mase, H., Kanamori, K., Ohwada, H.: Trust-aware recommender system incorporating review contents. Int. J. Mach. Learn. Comput. **4**(2) (2014)

5. Nguyen, M.H., Tran, D.Q.: A combination trust model for multi-agent systems. Int. J. Innov. Comput. Inf. Control **9**(6), 2405–2420 (2013)
6. Podobnik, V., Striga, D., Jandras, A., Lovrek, I.: How to calculate trust between social network users? In: 2012 20th International Conference on Software, Telecommunications and Computer Networks (SoftCOM), pp. 1–6. IEEE (2012)
7. Sherchan, W., Nepal, S., Paris, C.: A survey of trust in social networks. ACM Comput. Surv. **45**(4), 47:1–47:33 (2013)
8. Song, Y., Zhang, L., Giles, C.L.: Automatic tag recommendation algorithms for social recommender systems. ACM Trans. Web **5**(1), 4:1–4:31 (2011)
9. Tran, D.Q., Nguyen, M.H.: Classes of trust functions for distributed intelligent computing. Southeast Asian J. Sci. **1**(2), 148–154 (2012)
10. Wang, X., Liu, H., Fan, W.: Connecting users with similar interests via tag network inference. In: Proceedings of the 20th ACM International Conference on Information and Knowledge Management, CIKM 2011, pp. 1019–1024. ACM, New York (2011)
11. Zhang, L., Fang, H., Ng, W.K., Zhang, J.: Intrank: interaction ranking-based trustworthy friend recommendation. In: 2011 IEEE 10th International Conference on Trust, Security and Privacy in Computing and Communications, pp. 266–273, November 2011

Integrating Multisignature Scheme into the Group Signature Protocol

Hung Dao Tuan[1(✉)], Hieu Minh Nguyen[2], Cong Manh Tran[3], Hai Nam Nguyen[2], and Moldovyan Nikolay Adreevich[4]

[1] National Laboratory of Information Security, Ha Noi, Viet Nam
daotuanhung@gmail.com
[2] Academy of Cryptography Techniques, Ha Noi, Viet Nam
[3] Le Quy Don Technical University, Ha Noi, Viet Nam
[4] Saint-Petersburg Electronical University, St. Petersburg, Russia

Abstract. This paper proposes two new variants of group signature protocols with and without distinguished signing authorities based on the multisignature signature scheme to reduce significantly the signature length and masking signers public keys. The proposed protocols do not include a secret sharing and knowledge proving procedure. Thus, these protocols allow a flexible modification of the group structure by the group manager. Compared to the known group signature protocols, our protocols are designed based on integrating the multisignature scheme into the group signature protocol.

Keywords: Digital signature · Distinguished signing responsibilities · Discrete logarithm problem · Group signature · Multisignature · Public key

1 Introduction

Nowadays, the electronic transactions over the internet are performed not only between two individuals but also between groups of people or different organizations. Therefore, the informative authentication is to identify the information of a group of people or an organization. To solve this problem, several schemes such as multisignature [1], group signature [2], ring signature [3] and traditional signatures [4,5] have been proposed. Recently, group-oriented digital signatures [1–3] are widely used.

The multisignature of an electronic document is generated when each of the designated users signs the document [6–8]. Thus, this signature can be considered as a representative of m individual signatures. Each of the signers is responsible for the content of document, which is signed. To verify a given multisignature, the public keys of all users who generated the signature need to be used. In addition, the multisignature schemes and individual signature schemes can use the same public key infrastructure (PKI). The implementation of these schemes allows changing a set of signers arbitrarily. The last two properties of the protocols represent the important characteristics of the multisignature.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_32

The group signature of an electronic message is generated by a group of signers, and one of them is a group manager [2,9–11]. To verify the group signature, a group public key needs to be used, and he/she can not reveal which particular group member signed the document. The group signature has the following important properties. Firstly, only group members can sign the document. Secondly, the group manager, who has both the document and the valid group signature, can reveal the group members signed the document. Finally, non-group members could not reveal the original signers, who generate the group signature. Group manager is a trusted party of the group signature protocol. He creates the secret parameters, which are used to generate the signature by the signers.

In this paper, two new group signature schemes are proposed to provide both the features of the multi-signature scheme and the group signature. The proposed schemes implement signing steps using the methods generating the signature of both the multi-signature and the group signature schemes.

This design provides a possibility to keep the individual public keys of group members in secret. In order to obtain this property, the group signature is generated in two steps. First, the pre-signature of group is provided. Second, the group manager computes the group signature using a given approval based on pre-signature. The pre-signature can be generated by any group member or a subset of group members. Since only the group manager can reveal the original signers of the document, the distribution of the secret values and using the PKI are not required in the proposed group signature protocol. Moreover, the set of signers involved in the group can be arbitrarily changed by the group manager, and the public key of the group manager is used as the group public key. As the result, the group signature protocols provide a processing procedure of document very close to daily practices such as a letter document preparation, signing and the approval.

Furthermore, in the proposed schemes, the group manager has an ability to define the individual of the group to sign a part of the document, and signing authorities group signature are distinguished. The proposed group signature protocols can significantly reduce the signature length up to 640 bits in the case of 80-bit security. The design of the proposed protocols is only based on the computational difficulty of the elliptic curve discrete logarithm problem.

The rest of the paper is organized as follows. In Sect. 2, the implementation of the protocols based on the elliptic curve over finite field is presented. The analysis results of the proposed signature group protocols are discussed in Sect. 3. Finally, the conclusion is given in Sect. 4.

2 The Proposed Group Signature Protocols

Currently, cryptographic protocols based on elliptic curves (EC) over finite field have been widely used, and digital signature schemes based on ECDLP have attracted many researches [12]. In this paper, we propose two protocols for the group digital signature based on the multisignature with the ECschnorr digital signature scheme [12].

2.1 Proposed Group Signature Protocol 1

In the proposed protocol, the group manager (GM) will generate text messages needed to sign M. Group manager will define a structure of the group of people who will sign (Group manager will decide the members involved in the signing process). Determining the members, who will sign the text message, is based on the calculation of random values z_i. The random values z_i are calculated based on three parameters such as one-way functions (P_i) of public keys of signing members, hash values (h) of the text message which needs to be signed, and a secret value which is known by only the Group manager.

The group manager will calculate the values $h = H(M)$ and $z_i = H(H(h||P_i||SE)||h||P_i)$. Then, the values (z_i, h) are sent to i-th signer. To generate group signature, the following procedures are performed.

Step 1: Key Generation Phase

1. Each member in the group of signers generates their private key as a random number k_i $(1 < k_i < q)$ and public key computed as the point $P_i = k_i G$, with $i = (1, 2, \ldots, m)$.

2. The group manager computes his public key as the point $P_{gm} = k_{gm} G$, where k_{gm} is his private key. The public group key P_{gm} is used to verify the group signature.

Step 2: Group Signature Generation Phase

1. The group manager computes EC point $U = z_1 P_1 + z_2 P_2 + \ldots + z_m P_m$, which serves as the first element of the group signature.

2. Each i-th group member generates a random number $1 < t_i < q$, computes the value $R_i = t_i G$ and sends R_i to the group manager.

3. The group manager generates the random number $1 < t_{gm} < q$. Then, EC points $R_{gm} = t_{gm} G, R = R_{gm} + R_1 + R_2 + \ldots + R_m$ are computed. Moreover, the second element of the group signature $e = H(M||x_R||x_U)$, where x_R and x_U are x-coordinates of EC points R and U, respectively, is also calculated. He sends the value e to the group members who initiated the protocol.

4. Each i-th signer computes their signature share $s_i = t_i + k_i z_i e \bmod q$ and sends it to the group manager.

5. The group manager verifies the correctness of each s_i by checking $R_i = s_i G - z_i e P_i$. If all signature shares s_i satisfy the verification procedure, then he computes his share $s_{gm} = t_{gm} + k_{gm} e \bmod q$ and the third element of the group signature $s = s_{gm} + s_1 + s_2 + \ldots + s_m$.

The group signature of the document M is a tuple (U, e, s), which consists of one EC point and two numbers.

Step 3: Group Signature Generation Phase

1. The verifier computes the hash of the document M as $h = H(M)$.

2. Using the group public key P_{gm} and the signature (U, e, s), he computes the EC point $\tilde{R} = sG - e(U + P_{gm})$.

3. He computes the value $\tilde{e} = H(M||x_{\tilde{R}}||x_U)$ and compares the values $\tilde{e}$ and e. If $\tilde{e} = e$ then the verifier concludes that the group signature is valid.

2.2 Proposed Group Signature Protocol 2 with Distinguished Signing Authorities

In this protocol, the group manager will generate text messages needed to sign $M = M_1||M_2||\ldots||M_m$ (we only consider cases in which the number of text messages needed to sign equals the number of persons who signed).

Group manager will define a structure of the group of people who will sign (Group manager will decide the members involved in the signing process). Determining the members, who will sign the text message, is based on the calculation of random values z_i. The random values z_i are calculated based on three parameters such as one-way functions (P_i) of public keys of signing members, hash values (h_i) of the text message which needs to be signed, and a secret value which is known by only the group manager.

The group manager will calculate the values $h_i = H(M_i)$ and $z_i = H(H(h_i||P_i||SE)||h_i||P_i)$. Then, the values (z_i, h_i) are sent to i-th signer. To generate group signature, the following procedures are performed.

Step 1: Key Generation Phase (similar to step 1 of protocol 1)

1. Each member in the group of signers generates their private key as a random number k_i $(1 < k_i < q)$ and public key computed as the point $P_i = k_i G$, with $i = (1, 2, \ldots, m)$.

2. The group manager computes his public key as the point $P_{gm} = k_{gm} G$, where k_{gm} is his private key. The public group key P_{gm} is used to verify the group signature.

Step 2: Group Signature Generation Phase

1. The group manager computes EC point $U = h_1 z_1 P_1 + h_2 z_2 P_2 + \ldots + h_m z_m P_m$, which serves as the first element of the group signature.

2. Each i-th group member generates a random number $1 < t_i < q$, computes the value $R_i = t_i G$ and sends R_i to the group manager.

3. The group manager generates the random number $1 < t_{gm} < q$. Then, EC points $R_{gm} = t_{gm} G, R = R_{gm} + R_1 + R_2 + \ldots + R_m$ are computed. Moreover, the second element of the group signature $e = H(M||x_R||x_U)$, where x_R and x_U are x-coordinates of EC points R and U, respectively, is also calculated. He sends the value e to the group members who initiated the protocol.

4. Each i-th signer computes their signature share $s_i = t_i + k_i z_i h_i e \bmod q$ and sends it to the group manager.

5. The group manager verifies the correctness of each s_i by checking $R_i = s_i G - z_i h_i e P_i$. If all signature shares s_i satisfy the verification procedure, then he computes his share $s_{gm} = t_{gm} + k_{gm} e \bmod q$ and the third element of the group signature $s = s_{gm} + s_1 + s_2 + \ldots + s_m$.

The group signature of the document M is a tuple (U, e, s), which consists of one EC point and two numbers.

Step 3: Group Signature Generation Phase

1. The verifier computes the hash of the document M as $h = H(M)$.

2. Using the group public key P_{gm} and the signature (U, e, s), he computes the EC point $\tilde{R} = sG - e(U + P_{gm})$.

3. He computes the value $\tilde{e} = H(M||x_{\tilde{R}}||x_U)$ and compares the values $\tilde{e}$ and e. If $\tilde{e} = e$ then the verifier concludes that the group signature is valid.

3 Analysis of the Proposed Group Signature Protocols

3.1 Correctness

Proof of the protocol 1 correctness is as follows:

$$\tilde{R} = (s_{gm} + \sum_{i=1}^{n} s_i)G - e(P_{gm} + \sum_{i=1}^{m} z_i P_i)$$

$$= (t_{gm} + k_{gm}e + \sum_{i=1}^{n}(t_i + k_i z_i e))G - e(k_{gm}G + \sum_{i=1}^{m} z_i k_i G)$$

$$= (t_{gm} + k_{gm}e + \sum_{i=1}^{n} t_i + \sum_{i=1}^{n} k_i z_i e - e k_{gm} - \sum_{i=1}^{n} k_i z_i e)G$$

$$= (t_{gm} + \sum_{i=1}^{n} t_i)G = t_{gm}G + (\sum_{i=1}^{n} t_i)G$$

$$= R_{gm} + \sum_{i=1}^{n} R_i = R \Rightarrow \tilde{e} = H(M||x_{\tilde{R}}||x_U) = H(M||x_R||x_U) = e$$

Proof of the protocol 2 correctness is as follows:

$$\tilde{R} = (s_{gm} + \sum_{i=1}^{n} s_i)G - e(P_{gm} + \sum_{i=1}^{m} h_i z_i P_i)$$

$$= (t_{gm} + k_{gm}e + \sum_{i=1}^{n}(t_i + k_i z_i h_i e))G - e(k_{gm}G + \sum_{i=1}^{m} z_i k_i h_i G)$$

$$= (t_{gm} + k_{gm}e + \sum_{i=1}^{n} t_i + \sum_{i=1}^{n} k_i z_i h_i e - e k_{gm} - \sum_{i=1}^{n} k_i z_i h_i e)G$$

$$= (t_{gm} + \sum_{i=1}^{n} t_i)G = t_{gm}G + (\sum_{i=1}^{n} t_i)G$$

$$= R_{gm} + \sum_{i=1}^{n} R_i = R \Rightarrow \tilde{e} = H(M||x_{\tilde{R}}||x_U) = H(M||x_R||x_U) = e$$

3.2 Signature Length

In the proposed protocols, In the case of 80-bit security, it is possible to use EC with parameter q having size approximately 160 bits. Digital signatures on document M consists of three components (U, e, s), the signature length is equal to 640 bits. Compared to the lengths of previous group signature schemes [9–11], the signature length of the proposed protocols are significantly reduced.

3.3 Mechanism of Masking Public Keys

In the proposed protocols, the first element U of the group signature is used to reveal the group members who generated the signature. Element U is computed based on randomizing values z_i, depending on the document to be signed, and public keys of the group members. Values z_i are computed using the secure one-way hash function and a secret value SE known only to the group manager. Namely, each individual masking parameter zi is computed as follows $z_i = H(H(h||P_i||SE)||h||P_i)$ (or $zi = H(H(h_i||P_i||SE)||h_i||P_i)$), where SE is the additional secret key of the group manager. This formula defines the individual masking parameter for each user since it depends on their public key. The given value z_i of i-th user is different for each document because the hash value from each document is included in the argument of the specified hash function H. On the other hand, no one except group manager could reveal the value of the masked public key of the given signer and document since parameter z_i also depends on the secret value SE known only to group manager.

3.4 Traceability

When revealing the identity of the signers, the group manager needs to provide the proof that the identified set of signers really produced the given group signature. To do this, the group manager has to present values $H(h||P_i||SE)$, h and P_i (or $H(h_i||P_i||SE)$, h_i and P_i). Therefore, he has to keep his secret parameter SE unrevealed. The revealing of the value $H(h||P_i||SE)$ (or $H(h_i||P_i||SE)$) does not give any information to a potential malefactor since this parameter is valid only for the given document and one particular public key P_i. It can not be used to identify the group members who signed another document.

3.5 Unforgeability

The group members, who signed the document, depend on the authorization of the group manager. The computation of the group signature includes calculating the member signatures and the signature of the group manager. Therefore, only the group manager has the ability to properly authenticate and generate the group signature for the signing group.

3.6 Anonymity

The given value z_i of i-th user is different for each document because the hash value from each document is included in the argument of the specified hash-function H. Only the group manager could reveal the value of the masked public key of the given signer and document since parameter z_i also depends on secret value SE which known only by the group manager. Therefore, the proposed mechanism provides the anonymity of the signers for any person including the original signers, who verifying the group signature.

3.7 Exculpability

In the proposed schemes, no group member (or even some members join to do together) can forge the signatures of other group members. Because, the calculation of the signature value of each individual depends on not only the private key of each member (P_i), but also the random value (z_i) which calculated for each group member by the group manager. Therefore, in order to forge the signature of a group member, they need to pass the signature check equation for each member of the group manager. That means they have to break the ECDLP problem.

3.8 Unlinkability

In the proposed schemes, identifying the two different signatures generated by one member (or group of members) is impossible, except for the group manager. This is because of that the public keys of one member (or group of members) in the group have been masked by the calculation procedure of the group manager, and the verification of signatures only use the public key of the group manager. Thus, no one can identify which member (or which group of members) who signed on a document.

4 Conclusion

In this paper, the implementation of the two new group signature protocols based on usage of the multisignature scheme has been proposed to allow decreasing in the signature length and increasing in the efficiency of the signature generation procedure. The masking mechanism is used to conceal the original group members and do not influence on the size of generated signature. Moreover, we proposed a mechanism for masking public keys of the original signers. This mechanism is based on computing the hash-function value from the argument depending on the document to be signed such as the public keys of the original signers and an additional secret value of the group manager. These proposed protocols possess the features of both multisignature and group signature, therefore, they ensure higher capacity of applications in practice compared with single multisignature or group signature scheme alone.

References

1. Itakura, K., Kakamura, K.: A public-key cryptosystem suitable for digital multisignatures. NEC Res. Dev. **71**, 1–8 (1983)
2. Chaum, D., Van Heyst, E.: Group signatures. In: Davies, D.W. (ed.) Advances in Cryptology – EUROCRYPT 1991. LNCS, vol. 547, pp. 257–265. Springer, Heidelberg (1991)
3. Rivest, R.L., Shamir, A., Tauman, Y.: How to leak a secret. In: Boyd, C. (ed.) ASIACRYPT 2001. LNCS, vol. 2248, pp. 552–565. Springer, Heidelberg (2001). doi:10.1007/3-540-45682-1_32

4. ElGamal, T.: A public key cryptosystem and a signature scheme based on discrete logarithms. In: Blakley, G.R., Chaum, D. (eds.) CRYPTO 1984. LNCS, vol. 196, pp. 10–18. Springer, Heidelberg (1985). doi:10.1007/3-540-39568-7_2
5. Rivest, R.L., Shamir, A., Adleman, L.: A method for obtaining digital signatures and public-key cryptosystems. Commun. ACM. **21**, 120–126 (1978)
6. Ham, L.: Digital multisignature with distinguished signing authorities. Electron. Lett. **35**, 294–295 (1999)
7. Lin, S., Wang, B., Li, Z.: Digital multisignature on the generalized conic curve over Zn. Comp. Secur. **28**, 100–104 (2009)
8. Farashb, M.S., Biswasc, G.P., Khand, M.K., Obaidate, M.S.: A pairing-free certificateless digital multisignature scheme using elliptic curve cryptography. Int. J. Comput. Math. 1–18 (2015)
9. Kiayias, A., Yung, M.: Group signatures with efficient concurrent join. In: Cramer, R. (ed.) EUROCRYPT 2005. LNCS, vol. 3494, pp. 198–214. Springer, Heidelberg (2005). doi:10.1007/11426639_12
10. Laguillaumie, F., Langlois, A., Libert, B., Stehlé, D.: Lattice-based group signatures with logarithmic signature size. In: Sako, K., Sarkar, P. (eds.) ASIACRYPT 2013. LNCS, vol. 8270, pp. 41–61. Springer, Heidelberg (2013). doi:10.1007/978-3-642-42045-0_3
11. Benoit, L., San, L., Fabrice, M., Khoa, N., Huaxiong, W.: Signature schemes with efficient protocols and dynamic group signatures from lattice assumptions. Cryptology ePrint Archive Report (2016)
12. Cohen, H., Frey, G., Avanzi, R., Doche, C., Lange, T.N., Vercauteren, F.: Handbook of Elliptic and Hyperelliptic Curve Cryptography. CRC Press, Boca Raton (2005)

Interestingnesslab: A Framework for Developing and Using Objective Interestingness Measures

Lan Phuong Phan[1(✉)], Nghia Quoc Phan[2], Ky Minh Nguyen[3], Hung Huu Huynh[4], Hiep Xuan Huynh[1], and Fabrice Guillet[5]

[1] Can Tho University, 3/2 Street, Ninh Kieu District, Can Tho City, Vietnam
{pplan,hxhiep}@ctu.edu.vn
[2] Tra Vinh University, No. 126 National Road 53, Ward 5, Tra Vinh City, Vietnam
nghiatvnt@gmail.com
[3] Can Tho University of Technology,
256 Nguyen Van Cu Street, Ninh Kieu District, Can Tho City, Vietnam
nmky@ctuet.edu.vn
[4] University of Science and Technology - University of Danang,
54 Nguyen Luong Bang Street, Lien Chieu District, Da Nang City, Vietnam
hhhung@dut.dun.vn
[5] Polytech Nantes, University of Nantes,
La Chantrerie rue Christian Pauc BP 50609, 44306 Nantes Cedex 3, France

Abstract. The objective interestingness measures play an important role in data mining because they are used for mining, filtering and ranking the patterns. However, there is no research that collects the measures fully as well as there is no tool that can: automatically calculate the interestingness values of the patterns by using those measures, and is the framework for rapidly developing the applications related to objective interestingness measures. This paper describes Interestingnesslab - a tool of the objective interestingness measures is developed in the R language. The main functions of the tool are: mining a set of association rules and presenting them by the cardinalities $(n, n_X, n_Y, n_{X\overline{Y}})$, calculating the interestingness value of an association rule according to 1 of 109 collected measures; calculating the interestingness values of the whole rule set in many measures selected by the user; discovering the tendencies in a data set and recommending the top N items to the user; and studying the specific behavior of a set of interestingness measures in the context of a specific dataset and in an exploratory data analysis perspective. With Interestingnesslab, the user can easily and quickly reuse its functions to develop his/her own applications.

Keywords: Objective interestingness measure · Interestingnesslab · Association rule · Recommender system

1 Introduction

The data mining process takes a data set as the input and generates the patterns (such as the association rules, the classification rules) as the output [5].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_33

In fact, the data mining process can create hundreds and thousands of patterns. The determination of the most useful patterns can be performed by using the interestingness measures to calculate the actual value of the patterns. The interestingness measures play an important role in mining data, regardless of the type of patterns. They can be used for: (1) - pruning the unattractive patterns during the data mining process to narrow the search space and thus improve the efficiency of mining. For example, a threshold on the support measure can be used to remove patterns with low support values during mining process; (2) - ranking the patterns according to their interestingness values; (3) - filtering the interesting patterns during the post-processing. If the interestingness measures are good, the cost of time and space in mining data will be reduced. Each interestingness measure characterizes a certain aspect of the data set, therefore, the users should select the appropriate measure meeting their needs, calculate the interesting values of the patterns in the selected measure, and then extract the useful patterns.

The interestingness measures can be divided into two categories: subjective measures and objective measures [2,11]. The subjective approach evaluates the patterns by using the target, the knowledge, and the belief of user. The objective approach uses the statistical characteristics of the patterns to evaluate the interestingness. The second approach is only based on the raw data and does not require knowledge on the users or the application. Most interestingness measures are the objective interestingness measures. The objective interestingness measures are studied, surveyed by many independent group of authors, and at different times, such as Tan et al. in 2004 [10], Geng et al. in 2006 [1], Huynh et al. in 2008 [1], Heravi et al. in 2010 [8]; Grissa et al. in 2012 [7]; and Tew et al. in 2014 [12]. However, these studies just focus on the measures suitable for their own research orientation, and often focus on the common measures. For example, Huynh at el ranked 40 objective interestingness measures with sensitivity values; and Tew et al. focused on an analysis of the rule-ranking behavior of 61 well-known interestingness measures.

Although, there are a lot of researches on the interestingness measures, there still exist some mistakes in some researches: (1) - cite the formula of some measures incorrectly (the formula is improper as it is presented in the original research); (2) - use a measure that is called by different names, but just mention one name and do not take a note (or do not know) the remaining names. The mistake or the omission could be repeated if the latter researches refer to and cite from the previous researches, thereby affecting the quality of research. Besides, at the present, there is no research that synthesizes the objective interestingness measures fully, especially the recently proposed measures. The synthesize of the objective interestingness measures will form a common, complete, and reliable reference system which enables the researchers to save a lot of time and effort when studying the association rules and the measures of data mining. Moreover, there is also no automatic tool that meets the following criteria: (1) - calculate the value of each association rule according to many objective interestingness measures; (2) - is created as a framework for quickly developing applications to

detect the useful patterns, and then these applications can be easily integrated to the tool; (3) - is developed in R, a language and environment for the statistical computing and graphics. From this analysis, we propose a tool, named Interestingnesslab, to aggregate objective interestingness measures fully as well as provide the main functions as the framework for developing and using the objective interestingness measures.

This paper is organized into 5 parts. The first part is the introduction. The second part presents interestingness values. The third part describes the overview architecture of Interestingnesslab. The fourth part is core functions of the tool. The last part concludes this paper.

2 Interestingness Values

2.1 Objective Interestingness Measures

The objective interestingness measures used for evaluating the quality of patterns (i.e. the association rules in this paper) use statistics derived from data to determine whether an association rule is interesting. As mentioned in Part 1, there is no research that synthesizes the objective interestingness measures fully.

To collect the objective interestingness measures effectively, some criteria are identified: (1) - be the objects of researches on the interestingness measures as well as be cited by many others papers, (2) - be published by the reliable sources such as IEEE, Springer, ACM, Science Direct; (3) - be independently studied by the groups of authors.

After being collected, analyzed and validated, there are 109 different objective interestingness measures (109 different formulae), and 21 groups in which each group consists of some measures called by different names but having the same formula (Appendix). Formulae will be used for calculating the interestingness value of the association rules.

2.2 Presentation of an Association Rule

Let $I = \{I_1, I_2, \ldots, I_m\}$ be the set of different attributes (items); $D = \{T_1, T_2, \ldots, T_n\}$ be a transaction database in which each record $T_i(i : 1 \ldots n)$ is a transaction, and T_i is a subset of items ($T_i \subseteq I$), an association rule [11] is denoted by $X \rightarrow Y$ where X is called antecedence, Y is called consequence, X and Y are the subsets of items, and $X \cap Y = \emptyset$. An association rule represents the implicative trend between the item sets.

The presentation of an association rule $X \rightarrow Y$ can be expressed by a set of 4 values n, n_X, n_Y, and $n_{X\overline{Y}}$. $\{n, n_X, n_Y, n_{X\overline{Y}}\}$ is called the cardinality of an association rule where n is the number of transactions; $n_X = card(X)(n_Y)$ is the number of transactions that have $X(Y)$; and the counter-example number $n_{X\overline{Y}} = card(X \cap \overline{Y})(\overline{Y}$ is the complementary set of $Y)$ is the number of transactions that have X but do not have Y(Fig. 2).

For example, the association rule $\{egg, meat\} \rightarrow \{beer\}$ mined from the data set in Fig. 1 is represented by the cardinality $\{5, 3, 3, 1\}$.

	Bread	Egg	Meat	Beer	Milk
T_1	1	1	0	0	0
T_2	1	0	1	1	1
T_3	0	1	1	1	0
T_4	1	1	1	1	0
T_5	1	1	1	0	0

Fig. 1. An example of a transaction database.

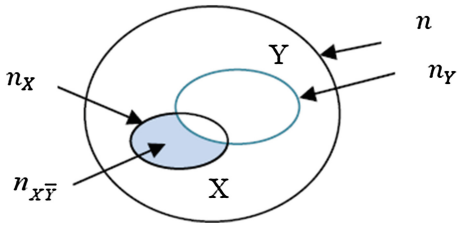


Fig. 2. The presentation of an association rule $X \rightarrow Y$.

2.3 Interestingness Value

The formula of an objective interestingness measure can be expressed by a function of 4 parameters n, n_X, n_Y, and $n_{X\overline{Y}}$: $m(X,Y) = f(n, n_X, n_Y, n_{X\overline{Y}})$. For example, the formula of the measure Support is $\frac{n_X - n_{X\overline{Y}}}{n}$. For 109 collected measures, their formulae are written in many different forms, such as the frequency, the number of transactions, etc. Therefore, for the convenient, all those formulae are converted to the functions of the cardinality n, n_X, n_Y, and $n_{X\overline{Y}}$.

The interestingness value (the quality) of an association rule $X \rightarrow Y$ in a measure is calculated by using the formula of that measure and the presentation of the rule $X \rightarrow Y$ (the cardinality $\{n, n_X, n_Y, n_{X\overline{Y}}\}$).

For example, if the association rule $\{egg, meat\} \rightarrow \{beer\}$ mined from the data set in Session 2.2 is represented by the cardinality $\{5, 3, 3, 1\}$, the interestingness value of this rule in the measure Support is $\frac{n_X - n_{X\overline{Y}}}{n} = \frac{3-1}{5} = 0.4$.

3 Architecture of Interestingnesslab

The overview architecture of Interestingnesslab is displayed as Fig. 3. The main components of this tool are: *cardinality*, *utility*, *application*, *interestingnessvalues*, and *interestingnessmeasures*.

The component *cardinality* is responsible for calculating the cardinalities of the rule set. It takes an association rule set generated by the Apriori algorithm, and a data set as the inputs; and generates the matrix *cardinality_matrix* as the output. Each row of *cardinality_matrix* includes the information: the ordinal

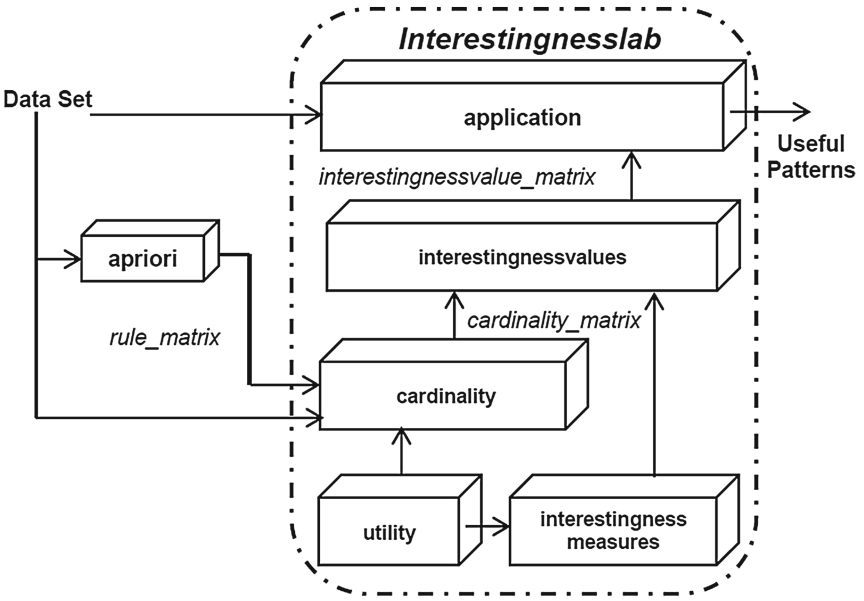


Fig. 3. The overview architecture of Interestingnesslab.

#	n	n_X	n_Y	$n_{X\overline{Y}}$	Presentation of rule
...	...	...	...	...	...
30	93	28	54	12	{Calculus} => { Probability and Statistics }
...	...	...	...	...	...

Fig. 4. An example of the matrix *cardinality_matrix*.

number (of a rule), $n, n_X, n_Y, n_{X\overline{Y}}$, the presentation of a rule in form $X \rightarrow Y$. Figure 4 shows an example of the matrix *cardinality_matrix*.

The component *utility* is a set of the utility functions that are used by the component *cardinality*.

The component *interestingnessvalues* is responsible for calculating the interestingness values of a rule set in the selected measures. This component takes *cardinality_matrix* as the input; generates *interestingnessvalue_matrix* as the output. Each row of *interestingnessvalue_matrix* consists of the information: the ordinal number (of a rule), $n, n_X, n_Y, n_{X\overline{Y}}$, the presentation of a rule in form $X \rightarrow Y$, the interestingness value of the first selected measure, the interestingness value of the second selected measure, etc. Figure 5 shows an example of the matrix *interestingnessvalue_matrix*.

The component *interestingnessmeasures* is a set of the functions where each function gets 4 parameters $n, n_X, n_Y, n_{X\overline{Y}}$ representing for an association rule; and returns the interestingness value of that association rule in a

#	n	n_X	n_Y	$n_{X\overline{Y}}$	Presentation of rule	Implication Intensity	Rule Interest
...	...	...	...	...	...	...	...
30	93	28	54	12	{Calculus} => { Probability and Statistics }	0.394548	-0.25806
...	...	...	...	...	...	...	...

Fig. 5. An example of the matrix *interestingnessvalue_matrix*.

specific measure. The function name is the measure name. The functions of *interestingnessmeasures* is used by the component *interestingnessvalues*.

The component *application* is an open component including the applications that are built by the users themselves as well as by four above components. At present, there are two applications already developed in this component: *ARQAT* and *ARbasedRS*. *ARQAT* (Association Rule Quality Analysis Tool) studies the specific behavior of a set of the interestingness measures in the context of a specific dataset and in an exploratory data analysis perspective. This tool implements 14 graphical and complementary views structured on 5 levels of analysis: ruleset analysis, correlation and clustering analysis, best rules analysis, sensitivity analysis, and comparative analysis. *ARQAT* was first developed in Java by Huynh et al. [9]. To integrate this tool to Interestingnesslab, *ARQAT* is re-implemented in R. The detail description of this tool is presented in [9]. Therefore, this paper does not remind the functions of *ARQAT*. *ARbasedRS* (Association Rule based Recommender System) discovers tendencies in a data set, and recommends the top N items to a user.y

4 Some Core Functions of Interestingnesslab

4.1 Presenting a Rule Set in the Form $\{n, n_X, n_Y, n_{X\overline{Y}}\}$

An association rule $X \rightarrow Y$ can be represented by a cardinality $\{n, n_X, n_Y, n_{X\overline{Y}}\}$. The following algorithm shows how to calculate $n, n_X, n_Y, n_{X\overline{Y}}$ for each rule of the rule set.

The algorithm for calculating the cardinalities of a rule set

Input: a set of the association rules (*ruleset*), a transaction database (*data*).
Output: the cardinalities of a set of rules (*cardinality_matrix*).
Steps:
- Count the number of transactions n;
- Calculate n_X

  Transform the left hand side of *ruleset* into a matrix *lhsRules* in which $lhsRules[i, j] = TRUE$ if the item j is one element of the left hand side of rule i, and $lhsRules[i, j] = FALSE$ otherwise;
  Calculate the matrix (cross) product $lhsProduct = lhsRules * t(data)$;
  Count n_X of each rule i $n_X[i] = rowSum(lhsProduct[i])$;
- Calculate n_Y. The method for calculating n_Y is similar to the method for

calculating n_X, except that it uses the right hand side of *ruleset*;
- Calculate $n_{X\overline{Y}}$:
    Calculate n_{XY}. The method for calculating n_{XY} is similar to the method for calculating n_X, except that it uses both side of *ruleset*;
    Calculate $n_{X\overline{Y}}$ of each rule i: $n_{X\overline{Y}}[i] = n_X[i] - n_{XY}[i]$;
- Concatenate n, n_X, n_Y, and $n_{X\overline{Y}}$ with *ruleset* to create the matrix *cardinality_matrix*.

4.2 Calculating the Interestingness Value of an Association Rule

Using 109 formulae of the objective interestingness measures converted to $\{n, n_X, n_Y, n_{X\overline{Y}}\}$, 109 functions are implemented. Each function takes the values n, n_X, n_Y, and $n_{X\overline{Y}}$ representing for an association rule as the input, and returns an interestingness value of that rule as the output.

4.3 Calculating the Interestingness Value of a Rule Set

Instead of calculating the interestingness value of an association rule in a measure, this function allow a user to calculate the interestingness values of a rule set in selected measures.

The algorithm for calculating the interestingness values of a rule set

Input: the cardinalities of a rule set (*cardinality_matrix*), a list of selected measures (*measures*).

Output: the interestingness values of a rule set in selected measures (*interestingnessvalue_matrix*).

Steps:

- Set ruleNum = the number of rules (rows of *cardinality_matrix*);
- Set measureNum = the number of the selected measures;
- Calculate the interestingness values:

```
for (i=1; i≤ruleNum; i++)
    Access the cardinality of rule i:
    for (j=1; j≤measureNum; j++)
        Calculate the interestingness value of rule i in measure
        j with name m: v = m(n, n_X, n_Y, n_{XȲ})
        Write the value v to the matrix value: values[i][j] = v
    }
}
```

- Concatenate two matrices (*values* and *cardinality_matrix*) to return the matrix *interestingnessvalue_matrix*.

4.4 Discovering Tendencies and Recommending Top N Items

The application called Association Rule based Recommender System is implemented by using the above functions. This system is developed to discover the tendencies in a data set, and recommend the top items to a user.

The algorithm for discovering the tendencies and recommending the top N items

Input: a transaction database (*data*) including a subset of items T_a that u_a liked.

Output: the tendencies or the top N items.

Steps:

- Generate a set of rules (*ruleset*) by using the Apriori algorithm. The user can set the thresholds on two measures support and confidence;
- Calculate the cardinalities $\{n, n_X, n_Y, n_{X\overline{Y}}\}$ of the rule set;
- Select the objective interestingness measures that are suitable for the user's purpose;
- Calculate the interestingness value $c(X, Y)$ for each rule of the *ruleset* in the selected measure;
- Sort the rules with their interestingness values in the descending order;
- Select 2 options:

  Show the tendencies in a data set by using the threshold on the selected measure.

  Recommend to a user u_a the top N items that u_a can like:

  Find all matching rules $X \rightarrow Y$ for which $(X \subseteq T_a)$;

  Recommend N right hand sides (Y) of matching rules with the highest values.

5 Conclusion

This paper has collected and validated 109 objective interestingness measures, then converted their formula to the unified format (the cardinality $\{n, n_X, n_Y, n_{X\overline{Y}}\}$). The list of these measures can be regarded as a complete, systematic, and reliable reference source. Besides, the tool of the objective interestingness measures, named Interestingnesslab, has been developed with the main functions: presenting an association rule set by the cardinalities; calculating the interestingness values of a rule in a specific measure; calculating the interestingness values of the rule set in measures selected by the user; building an application to detect the tendencies in a data set and to recommend the top N items to a user; and studying the specific behavior of a set of the interestingness measures in the context of a specific dataset and in an exploratory data analysis perspective. Interestingnesslab is implemented in the R language, and is an open source package. Therefore, the users can fully reuse the core functions to develop and use their own applications.

Appendix

21 groups of measures, each group includes the measures called by the different names but having the same formula.

#	Group of measures	#	Group of measures
1	Accuracy, Causal Support	2	Added Value, Pavillon, Centered Confidence
3	Bayes Factor, Odd Multiplier	4	Correlation Coefficient, Phi-Coefficient, Pearson's Correlation Coefficient, Linear-Correlation, Newrelevancy
5	Cosine, Ochia, IS Measure	6	Descriptive Confirmed-Confidence, Lerman Similarity Index
7	Dice Index, Czekanowski Dice, F-Measure Examples and Contra-Examples	8	Directed Contribution to Chi square, Lerman Similarity Index
9	Rate, Example and Contra-Example Rate, Encountered Rate	10	Gray and Orlowska's Interestingness Weighting Dependency, I-Measure
11	Indice Probabilistc d'Ecart d'Equilibre, Probabilistic Measure of Deviation from Equilibrium(IPEE)	12	Jaccard, Coherence
13	Kappa Coefficient, Cohen	14	Kulczynski 1, Agreement-Disagreement Index
15	Lift, Interest	16	Loevinger, Certainty Factor, Satisfaction
17	Mutual Information, 2-way Support Variation	18	Normalized Difference, Match
19	Piatetsky-Shapiro, Pearl, Leverage 2, Carnap, Novelty	20	Relative Risk, Class Correlation Ratio
21	Specificity 1, Negative Reliability		

References

1. Huynh, H.X., Guillet, F., Le, T.Q., Briand, H.: Ranking objective interestingness measures with sensitivity values. VNU J. Sci. Nat. Sci. Technol. **24**, 122–132 (2008)
2. McGarry, K.: A survey of interestingness measures for knowledge discovery. Knowl. Eng. Rev. J. **20**(1), 39–61 (2005)
3. Guillet, F., Hamilton, H.Z.: Quality Measures in Data Mining. Series in Computational Intelligence, vol. 43. Springer, Heidelberg (2007)
4. Tan, P.-N., Steinbach, M., Kumar, V.: Introduction to Data Mining. Pearson (2006)
5. Geng, L., Hamilton, H.J.: Interestingness measures for data mining: a survey. ACM Comput. Surv. **38**(3) (2006). Article 9
6. Guillaume, S., Grissa, D., Nguifo, E.M.: Categorization of interestingness measures for knowledge extraction. CoRR abs/1206.6741 (2012)
7. Grissa, D., Guillaume, S., Nguifo, E.M.: Combining Clustering techniques and FCA to characterize Interestingness Measures, Research Report LIMOS/RR-12-05 (2012)

8. Heravi, M.J., Zaïane, O.R.: A study on interestingness measures for associative classifiers. In: Proceedings of the 2010 ACM Symposium on Applied Computing, SAC 2010, pp. 1039–1046 (2010)
9. Huynh, X.H., Guillet, F., Briand, H.: ARQAT: an exploratory analysis tool for interestingness measures. In: Proceedings of the 11th International Symposium on Applied Stochastic Model and Data Analysis, ASMDA 2005, Brest, France, pp. 334–344 (2005)
10. Silberschatz, A., Tuzhilin, A.: What makes patterns interesting in knowledge discovery systems. IEEE Trans. Knowl. Data Eng. **8**(6), 970–974 (1996)
11. Tan, P.-N., Kumar, V., Srivastava, J.: Selecting the right objective measure for association analysis. Inf. Syst. **29**(4), 293–313 (2004)
12. Tew, C., Giraud-Carrier, C., Tanner, K., Burton, S.: Behavior-based clustering and analysis of interestingness measures for association rule mining. Data Mining Knowl. Discov. **28**(4), 1004–1045 (2014). Springer, US

Inverted Pendulum Control Using Fuzzy Reasoning Method Based on Hedge Algebras by Approach to Semantic Quantifying Adjustment of Linguistic Value

Nguyen Duy Minh(✉), Do Thi Mai, and Nguyen Thi Thu Hien

University of Information and Communication Technology,
Thai Nguyen University, Thai Nguyen, Vietnam
ndminh@ictu.edu.vn

Abstract. Fuzzy control using hedge - algebras is used in the control problem with the linguistic value of k finite depth. With semantics quantifying adjustment of the linguistic value that still preserve their order. This paper proposes the fuzzy reasoning method based on hedge algebras (HAs) by approach to semantic quantifying adjustment of linguistic value, the semantic quantifying adjustable parameters are determined by genetic algorithm. This method is applied to the problem of inverted pendulum control of Ross, the simulation results have confirmed that the given method is correct and effective.

Keywords: Inverted pendulum · Hedge algebra · Reasoning method · Semantic quantifying · Fuzzy control

1 Introduction

In fuzzy control. Dependency between physical variable is expressed not by mathematical equation, but by fuzzy model of the following form:

$$\begin{array}{c} If\ X_1 = A_{11}\ and ... and\ X_m = A_{1m}\ then\ Y = B_1 \\ If\ X_1 = A_{21}\ and ... and\ X_m = A_{2m}\ then\ Y = B_2 \\ ..... \\ If\ X_1 = A_{n1}\ and ... and\ X_m = A_{nm}\ then\ Y = B_n \end{array} \quad (1)$$

where $A_{i,j}, j = 1, ..., m$, and $B_i, i = 1, ..., n$ are verbal descriptions of the linguistic variable X_j and Y, respectively [8] and the notation, $X_j = A_{ij}$ or $(Y = B_i)$ is an abbreviation of the sentence X_j is A_i or Y is B_i. Therefore, the kernel of fuzzy controller is the fuzzy multiple conditional reasoning (FMCR) which says given input $X_j = A_{0j}, j = 1, ..., m$, of fuzzy model (1), find an output $Y = B_0$. The fuzzy model (1) in the fuzzy control is also called fuzzy associative memory (FAM) which represents an expert knowledge of an application domain.

In a number of recent researches [5,7], authors have developed the fuzzy reasoning method based on HAs and were shown good results.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_34

However, reasoning method based on hedge – algebras for solving real problems with linguistic value of k finite depth. With limit the depth of linguistic value, we can completely adjust quantitative semantics of linguistic values that still preserve their order. The goal of this paper is determining the adjustable quantifying semantic parameters of fuzzy reasoning method based on HAs and application to inverted pendulum control.

For realization of this goal, in Sects. 2 and 3 of the article we recall some results, which have a connection with the quantification of linguistic value, the problems of quantifying semantic adjustment of linguistic values, in Sects. 4 and 5 we represent researches into the fuzzy reasoning method based on HAs by approach to semantic quantifying adjustment and application.

2 Fuzziness Measure of Linguistic Value and Semantically Quantifying Mapping

Let HAs $\underline{AX^*} = (\underline{X}^*, G, H, \sigma, \phi, \leq)$ be a complete linear hedge algebras (CLinHA), where $\underline{X}^*$ is base set, $G = (0, c^+, W, c^+, 1)$ is set of generated elements, H is set of positive and negative hedges, $\leq$ is complete ordering relation on $\underline{X}^*$, σ and ϕ are two extended mathematical operations so that with all of $x \in \underline{X}^*$, ϕx, σx respectively lower right threshold and upper right threshold on $\underline{X}^*$ of set $H(x)$, which is the set of all elements generated by hedges on H, $H = H^- \cup H^+$, and assuming that $H^- = \{h_{-1}, ..., h_{-q}\}$, with $h_{-1} < h_{-2} < ... < h_{-q}$, and $H^+ = \{h_1, ..., h_p\}$, with $h_1 < h_2 < ... < h_q$ where by convention we have $h_0 = I$, unit operator on $\underline{X}^*$. On the supposition that $\underline{AX^*}$ is free HAs, that is $\forall x \in H(G), \forall h \in H, hx \neq x$.

Definition 1 [2,3]. *Function $fm : X^* \to \{-1, 0, 1\}$ is said to be a fuzziness measure (FM, for short) of terms in X, if:*

(F1) fm is a complete measure on X^*, so that $fm(c^-) + fm(c^+) = 1$ and $\forall u \in X^*, \sum_{h \in H} fm(hu) = fm(u)$;

(F2) If x is a exact concept, so that $H(x) = \{x\}$ then $fm(x) = 0$.

Especially, $fm(0) = fm(W) = fm(1) = 0$;

(F3) $\forall x, y \in X^*, \forall h \in H$, we have $\frac{fm(hx)}{fm(x)} = \frac{fm(hy)}{fm(y)}$, that is this proportion does not depend on specific elements and, hence, it is called fuzziness measure of the hedge h and denoted by $\mu(h)$.

Proposition 1 [2,3]. *For each fuzziness measure fm on X the following statements hold:*

(1) $fm(hx) = \mu(h) fm(x), \forall x \in X^*$
(2) $fm(c^-) + fm(c^+) = 1$;
(3) $\sum_{i=-q, i\neq 0}^{p} fm(h_i c) = fm(c)$ *with* $c \in \{c^-, c^+\}$;
(4) $\sum_{i=-q, i\neq 0}^{p} fm(h_i x) = fm(x)$;
(5) $\sum_{i=-1}^{-q} \mu(h_i) = \alpha$ *and* $\sum_{i=1}^{p} \mu(h_i) = \beta$, *with* $\alpha, \beta > 0$ *and* $\alpha + \beta = 1$.

Definition 2 [2]. *A function* $Sign : X \to \{-1, 0, 1\}$ *is a mapping which is defined recursively as follows, for* $h, h' \in H$ *and* $c \in \{c^-, c^+\}$*:*

(a) $Sign(c^-) = -1, Sign\,(c^+) = +1$;
(b) $Sign\,(hc) = -Sign\,(c)$ *if* $hc \neq c$ *and* h *is negative with respect to (w.r.t)* c;
(c) $Sign\,(hc) = Sign\,(c)$ *if* $hc \neq c$ *and* h *is positive w.r.t* c;
(d) $Sign\,(h'hx) = -Sign\,(hx)$ *if* $h'hx \neq hx$ *and* h' *is negative w.r.t* h;
(e) $Sign\,(h'hx) = Sign\,(hx)$ *if* $h'hx \neq hx$ *and* h' *is positive w.r.t* h;
(f) $Sign\,(h'hx) = 0$ *if* $h'hx = hx$.

Definition 3 [3]. *Let* fm *be a fuzziness measure on* X. *A mapping* $\nu : X \to \{0, 1\}$, *which is induced by* fm *on* X, *is defined as follows:*

(1) $\nu\,(W) = q = fm(c^-), \nu(c^-) = \theta - afm(c^-), \nu\,(c^+) = \theta + afm\,(c^+)$;
(2) $\nu(h_j x) = \nu(x) + Sign(h_j x)(\sum_{i=sign(j)}^{j} fm(h_i x) - \omega(h_j x) fm(h_j x))$, *where* $j \in [-q^\wedge p] = \{j : -q \leq j \leq p, j \neq 0\} = j \in [-q^\wedge p]$ *and* $\omega(h_j x) = \frac{1}{2}[1 + Sign(h_j x) Sign(h_p h_j x)(\beta - \alpha)] \in \{\alpha, \beta\}$;
(3) $\nu(\phi c^-) = 0, \nu(\sigma c^-) = \theta = \nu(\phi c^+), \nu(\sigma c^+) = 1$ *and for* $j \in [-q^\wedge p]$,
$\nu(\phi h_j x) = \nu\,(x) + Sign(h_j x) \left(\sum_{i=Sign(j)}^{j-Sign(j)} \mu(h_i) fm(x)\right)$
$-\frac{1}{2}\,(1 - Sign(h_j x))\,\mu(h_j) fm(x)$,
$\nu(\sigma h_j x) = \nu\,(x) + Sign(h_j x) \left(\sum_{i=Sign(j)}^{j-Sign(j)} \mu(h_i) fm(x)\right)$
$+\frac{1}{2}\,(1 + Sign(h_j x))\,\mu(h_j) fm(x)$.

3 Problem of Semantically Quantifying Adjustment in Linguistic Value

Definition 4 [6]. *Real number* $\varepsilon, 0 < \varepsilon < 1$, *is called threshold of semantically quantifying adjustment in linguistic value on* X_k *if with all of* $x, y \in X_k$ *satisfies* $x < y$ *brings about* $\nu\,(x) \pm \varepsilon < \nu\,(y) \pm \varepsilon$.

Theorem 1 [6]. *Let* AX^* *be a linear, complete and free HAs, always exist threshold of semantically quantifying adjustment in linguistic value on* X_k *with* k *is any positive number.*

Definition 5 [6]. *Let* AX^* *be a HAs with* $\forall k \geq 1$ *threshold of semantically quantifying adjustment in linguistic value on* X_k *is determined as:*

$$e_k = min\left\{\frac{\alpha fm(x)}{2}, \frac{\beta fm(x)}{2}\right\}, |x \in X_k\}$$

.

4 Fuzzy Reasoning Method Based on HAs by Approach to Semantically Quantifying Adjustment

Fuzziness reasoning method based on HAs depends on choosing parameters for HAs of linguistic variables, normally use intuitive method to choosing these parameters also is reasonable. Thus, we can intuitively fix parameters of HAs, however to create flexibility of method, we can adjust in semantically quantifying of linguistic value by entering some adjustable parameters, which must be less than threshold of adjustment as is shown in Definition 5.

One problem is finding out: how to find adjustable quantifying parameters? Return to the reasoning method using HAs, which is mentioned in Sect. 3. It is assumed that parameters of HAs and linguistic variables, respectively, in the model FAM (Fuzzy Associative Memory) are preselected and model FAM has m linguistic variables (FAM of m – dimention), in this case we have m adjustable parameters, denoted by PAR and then HAR depends on the set of these parameters.

This method consists of the following steps:

Step 1: Construct HAs AX_j for linguistic variables X_j and AY for linguistic variables Y. Assuming that HAs AX_j, and $A_{ij}, i = 1, \ldots, n; j = 1, \ldots, m$ are linguistic values of linguistic variables X_j, the set of semantically quantifying values of variables X_j is

$$(\nu_{Xj}(A_{\mathbf{1}j}), \nu_{Xj}(A_{\mathbf{2}j}), \ldots, \nu_{Xj}(A_{ij}));$$

$HA\ AY$, and $B_i, i = 1, \ldots, n$ are linguistic values of linguistic variables Y, the set of semantically quantifying values of variables is $(\nu_Y(B_{\mathbf{1}}), \nu_Y(B_{\mathbf{2}}), \ldots, \nu_Y(B_i))$.

Step 2: Using semantically quantifying mapping (SAM) to determine model SAM. Construct adjustable model quantifying associative memory, so that we enter the adjustable semantically quantifying parameters $((\nu_{Xj}(A_{\mathrm{ij}}) + \delta_{\mathrm{ij}}), (\nu_Y(B_i) + \delta_i))$, where δ_{ij} is adjustable semantically quantifying parameter of variable X_j, δ_I is adjustable semantically quantifying parameter of variable Y, satisfy condition $|\delta_{ij}| < \varepsilon_{X_j}, |\delta_i| < \varepsilon_Y$ with $\varepsilon_{X_j}, \varepsilon_Y$ are adjustable thresholds of X_j and Y.

Step 3: Construct reason by linear interpolative method: we choose a weighted averaging operator with weights $w_1, ..., w_2$ [8] to aggregate the first m components of each point of $C_{R,m+1}$ and obtain a real curve $C_{R,2}$ in $[0,1] \times [0,1]$. The algorithm establishes a real curve $C_{R,2}$ defined by an model SAM of adjustable semantically quantifying parameters (PAR, for short), is called SAM(PAR).

Step 4: Apply the linear classical interpolative method to $C_{R,2}$ to compute the output $\nu_Y(B_0)$ corresponding to the given input.

For convenience in presenting: Adjustable Model of Semantically Quantifying Associative Memory is denoted by SAM(PAR); Fuzzy Reasoning Method using HAs by approach to adjustment in semantically quantifying value with threshold, is denoted by OpHAR.

It is assumed that SAM has m – input linguistic variables $X_j, j = 1, ..., m$, and output linguistic variables Y, when for each linguistic variable X_j

(fixed $j = 1, \ldots, m$), we have n - adjustable semantically quantifying parameters of linguistic variables Y and in general is presenting as follow:

- Adjustable semantically quantifying parameters of linguistic variables X_j are $((\delta_{11}, \delta_{21}, \ldots, \delta_{n1}), (\delta_{12}, \delta_{22}, .., \delta_{n2}), \ldots, (\delta_{1m}, \delta_{2m}, \ldots, \delta_{nm}))$.
- Adjustable semantically quantifying parameters of linguistic variables Y are $(\delta_1, \delta_2, ..., \delta_n)$.

The set of adjustable semantically quantifying parameters of linguistic variables is:

$$PAR = ((\delta_{11}, \delta_{21}, \ldots, \delta_{n1}), (\delta_{12}, \delta_{22}, .., \delta_{n2}), \ldots, (\delta_{1m}, \delta_{2m}, .., \delta_{nm}), (\delta_1, \delta_2, .., \delta_n)). \quad (2)$$

Optimization of the parameters in the set PAR: To find the parameters PAR, we construct the following optimization algorithm using a GA, with goal function g will get the minimum.

$$g = f((\delta_{11}, \delta_{21}, \ldots, \delta_{n1}), (\delta_{12}, \delta_{22}, .., \delta_{n2}), \ldots, (\delta_{1m}, \delta_{2m}, .., \delta_{nm}), (\delta_1, \delta_2, .., \delta_n)) \rightarrow \min$$

with constrain condition:

$$\begin{array}{l} |\delta_{ij}| < e_{Xj}; i = 1, \ldots, n; j = 1, \ldots, m. \\ |\delta_i| < e_Y; i = 1, \ldots, n. \end{array} \quad (3)$$

5 Control of the Inverted Pendulum Using Fuzzy Reasoning Method Based on HAs by Approach to Adjustment in Quantifying Semantics

Describe? the inverted pendulum problem:

The task of inverted pendulum control of Ross [5], is the canonical task with the non – linear investigative system (Fig. 1). The goal of the control is putting the inverted pendulum to the balanced position. The differential equation of the inverted pendulum system is described as follow:

$$-ml^2 d^2\psi/dt^2 + mlg \sin\psi = u(t) \quad (4)$$

where m is the mass of the pole located at the tip point of the pendulum; l is the length of the pendulum; ψ is the deviation angle from vertical in the clockwise direction; $u(t)$ is the control action; t is the time; g is the gravitational acceleration constant.

It is assumed that $x_1 = \psi$ and $x_2 = \frac{d\psi}{dt}$ are the state variables. In the case that deviation angle is very small then $\sin(\psi) = \psi, \sin(\psi)$ is measured in radian.

Choosing the length of the pendulum $l = g$ and mass $m = 180/(\pi.g^2)$, by linearization the initial system as in [5], we have system, which consists two linear discrete equations without measure.

$$\begin{array}{l} x_1(k+1) = x_1(k) + x_2(k) \\ x_2(k+1) = x_1(k) + x_2(k) - u(k) \end{array} \quad (5)$$

It is assumed that the universes of discourse of the two variable are $-2^{o} \leq x_1(k) \leq 2^{o}$, $-5\,\text{dps} \leq x_1(k) \leq 5\,\text{dps}$ and of the control action is $-16\,\text{mA} \leq u(k) \leq 16\,\text{mA}$ which are the same as given in [5].

The required task: Find out the value of the control signal u based on control rules of the inverted pendulum, so that the inverted pendulum always is kept in the clockwise direction, this position is called balance, so that $x_1 = 0$ and $x_2 = 0$.

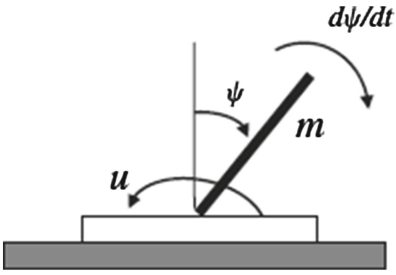


Fig. 1. Inverted pendulum control problem

The error of the system:

$$e(k) = \sqrt{x_1^2(k) + x_2^2(k)} \tag{6}$$

Using fuzzy reasoning method based on HAs by approach to adjustment in quantifying semantics to control the inverted pendulum: According to the document [5], system of control rules the inverted pendulum consists the linguistic labels of variables X_1, X_2 are P (Positive), N (Negative), Z (Zero); u is P (Positive), N (Negative), Z (Zero), PB (Positive Big), NB (Negative Big). The control rules are designed for fuzzy control process shown in the Table 1.

Table 1. Model FAM for inverted pendulum system

X_1/X_2	P	Z	N
P	PB	P	Z
Z	P	Z	N
N	Z	N	NB

Step 1: Construct parameters of HAs:
The linguistic labels are transform as follow:

Table 2. Model SAM for inverted pendulum system

X_1/X_2	L: 0.75	W: 0.5	S: 0.25
L: 0.75	ML: 0.8	PL: 0.7	W: 0.5
W: 0.5	PL: 0.7	P: 0.5	PS: 0.3
S: 0.25	W: 0.5	PS: 0.3	MS: 0.2

- Linguistic variables X_1, X_2 have linguistic values $N \rightarrow Small$, $P \rightarrow Large$, $Z \rightarrow W$, and quantifying value of linguistic values $\nu_{X_1}(Large) = 0.75$, $\nu_{X_1}(W) = 0.5$, $\nu_{X_1}(Small) = 0.25$, $\nu_{X_2}(Large) = 0.75$, $\nu_{X_2}(W) = 0.5$, $\nu_{X_2}(Small) = 0.25$
- Linguistic variables u have linguistic values $NB \rightarrow MoreSmall$, $N \rightarrow PossiblySmall$, $Z \rightarrow W$, $P \rightarrow PossiblyLarge$, $PB \rightarrow MoreLarge$, and quantifying value of linguistic values $\nu_u(MoreLarge) = 0.875$, $\nu_u(PossiblyLarge) = 0.7$, $\nu_u(W) = 0.5$, $\nu_u(PossiblySmall) = 0.3$, $\nu_u(MoreSmall) = 0.2$.

Result of semantically quantifying value of linguistic variable X_1, X_2 and u as in the Table 2.

Step 2: From model SAM in the Table 2, construct linguistic variables X_1, X_2 and u as follow:

- The linguistic variable X_1 has linguistic values $N \rightarrow Small$, $P \rightarrow Large$, $Z \rightarrow W$, there are linguistic values of the depth $k = 1$, according to theorem 1, adjustable threshold is 0.125 and quantifying values of linguistic values is determined as follow:

$$\nu_{X1}(Large) = 0.75 + \delta_{X11}; \nu_{X1}(W) = 0.5 + \delta_{X12}; \nu_{X1}(Small) = 0.25 + \delta_{X13}$$

- Similar to X_1, the linguistic variable X_2 has linguistic values $N \rightarrow Small$, $P \rightarrow Large$, $Z \rightarrow W$, there are linguistic values of the depth $k = 1$, adjustable threshold is 0.125 and quantifying values of linguistic values are:

$$\nu_{X1}(Large) = 0.75 + \delta_{X11}; \nu_{X1}(W) = 0.5 + \delta_{X12}; \nu_{X1}(Small) = 0.25 + \delta_{X13}$$

The linguistic variable u has linguistic values $NB \rightarrow MoreSmall$, $N \rightarrow PossiblySmall$, $Z \rightarrow W$, $P \rightarrow PossiblyLarge$, $PB \rightarrow MoreLarge$, there are linguistic values of the depth $k = 2$, adjustable threshold is 0.0625 and quantifying values of linguistic values is determined as follow:

$$\nu_u(MoreLarge) = 0.875 + \delta_{u7}; \nu_u(PossiblyLarge) = 0.7 + \delta_{u8}; \nu_u(W) = 0.5 + \delta_u;$$
$$\nu_u(PossiblySmall) = 0.3 + \delta_{u10}; \nu_u(MoreSmall) = 0.2 + \delta_{u11}$$

$PAR = \{\delta_{X1i}, i = 1, ..., 3; \delta_{X2i}, i = 4, ..., 6; \delta_{ui}, i = 7, ..., 11\}$, with conditions:

$$|\delta_{X_1 i}| < 0.125, i = 1, ..., 3 \text{ for variable } X_1$$
$$|\delta_{X_2 i}| < 0.125, i = 4, ..., 6 \text{ for variable } X_2$$
$$|\delta_{ui}| < 0.125, i = 7, ..., 11 \text{ for variable u}$$

Table 3. Model SAM(PAR) of inverted pendulum system

X_{1s}/X_{2s}	$0.75+\delta_{X24}$	$0.5+\delta_{X25}$	$0.25+\delta_{X26}$
$0.75+\delta_{X11}$	$0.8+\delta_{u7}$	$0.7+\delta_{u8}$	$0.5+\delta_{u9}$
$0.5+\delta_{X12}$	$0.7+\delta_{u8}$	$0.5+\delta_{u9}$	$0.3+\delta_{u10}$
$0.25+\delta_{X13}$	$0.5+\delta_{u9}$	$0.3+\delta_{u10}$	$0.2+\delta_{u11}$

Hence, model SAM(PAR) is defined as in the Table 3.

Step 3: Using linear interpolation and coordination with weight. Construct linear interpolation and coordination with weight $agg = X_{1s} * \mathrm{w}_1 + X_{2s} * \mathrm{w}_2$, with $w1 = 0.37066$ and $w2 = 0,62934$ [7].

Step 4: Determine the output.

Applying the method OpHAR to inverted pendulum system with initial values of states are concrete values $X_1(1) = 1.00$ and $X_2(1) = -4$, semantization and desemantization according to formulas 7 and 8:

$$\text{semantization}(x) = s_0 + \frac{s_1 - s_0}{x_1 - x_0}(x - x_0) \tag{7}$$

$$\text{desemantization}(s) = x_0 + \frac{x_1 - x_0}{s_1 - s_0}(s - s_0) \tag{8}$$

with

$$\begin{aligned} &s_0 = 0.75 + \delta_{X11}, s_1 = 0.25 + \delta_{X13} \text{ and } x_0 = 2, x_1 = -2 \text{ for } X_1 \\ &s_0 = 0.75 + \delta_{X24}, s_1 = 0.25 + \delta_{X26} \text{ and } x_0 = 5, x_1 = -5 \text{ for } X_2 \\ &s_0 = 0.8 + \delta_{u7}, s_1 = 0.2 + \delta_{u11} \text{ and } x_0 = 16, x_1 = -16 \text{ for } u \end{aligned}$$

The goal function e is defined as follow:

$$e(k) = \sum_{k=1}^{p} \sqrt{x_1^2(k) + x_2^2(k)} = \min \tag{9}$$

where, p is the number of control cycles. The parameters PAR of the optimal using genetic algorithms (GA) function e with number of generations: 200, recombination probability: 0.80; mutation probability: 0.05; number of individual per subpopulation: 40; number of individual: 10;

Through a number of times to run MATLAB simulations, we determine the model OpPAR and coordinative weights are:

$$\begin{aligned} OpPAR = \{&0.083456; 0.076369; -0.109360; -0.068060; -0.058040; \\ &-0.056329; -0.035618; 0.028409; 0.032319; 0.051259; 0.013991\} \end{aligned}$$

– The error of control is compared in the Table 4.

Look at the graph in the Fig. 2, we find out that the result of method OpHAR is better than others by methods in [7].

Table 4. The error of methods in inverted pendulum system

Method	Error of control
Fuzzy reasoning method using HAs (HAR) [1,7]	15.48957
Fuzzy reasoning method by optimization the parameters of HAs [7]	4.757936
Fuzzy reasoning method using HAs by approach to adjustment in semantics quantifying with threshold (OpHAR)	4.643746

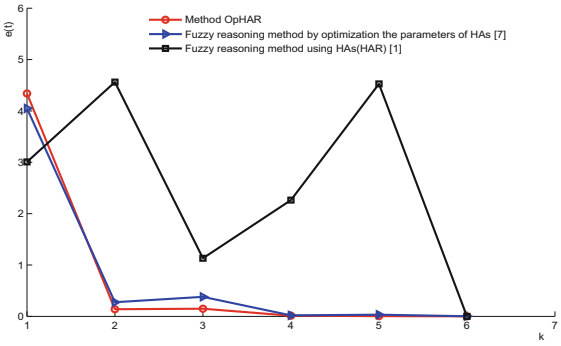


Fig. 2. Error graphs of inverted pendulum system

Comment: From the Table 4 and Fig. 2, we realize that total error was found through control cycles using reasoning method OpHAR is minimum in comparison with the total error when we use fuzzy reasoning method by optimization the parameters of HAs [7] and fuzzy reasoning method using HAs(HAR). Inverted pendulum has tendency to balance position and goal function accept minimal value through 2 control cycles by using method Op(HAR).

6 Conclusion

Semantics quantifying adjustment of linguistic value allows us choose parameters of HAs by intuition. It makes the reasoning method is much closer to human thinking. On the other side, with this approach we just to construct one HAs for all of linguistic variables in the task, makes the method become easier and more simple to use.

Based on the mathematic, adjustment in semantics quantifying of linguistic values in HAs [6], this paper presented fuzzy reasoning method using HAs by approach to adjustment in semantics quantifying of linguistic values in inverted pendulum control.

Result is applied in control task of inverted pendulum shows that method of approach to adjustment in semantics quantifying of linguistic values, which is presented in this paper is accurate, effective, and promise applying to continuous tasks.

References

1. Lan, N.V., Hung, C.V., Minh, D.N., Xuan, N.D.: Control the inverted pendulum based on HAs. J. Sci. Technol. **43**(06), 9–18 (2005)
2. Ho, N.C., Lan, V.N., Viet, X.L.: Fuzziness measure on complete hedge algebras and quantifying semantics of terms in linear hedge algebras. Fuzzy Sets Syst. **154**(4), 452–471 (2007)
3. Ho, C.N., Long, V.N.: The mathematical basic of fuzziness measure of language information. J. Comput. Sci. Cybern. **20**(01), 64–72 (2004)
4. Ho, N.C., Wechler, W.: Hedge algebra: an algebraic approach to structures of sets of linguistic truth values. Fuzzy Sets Syst. **35**, 281–293 (1990)
5. Ross, T.J.: Fuzzy logic with Engineering Applications, International edn. McGraw-Hill, New York (2004)
6. Minh, D.N.: Adjustment in semantics quantifying of linguistic value in HAs and application. J. Sci. Technol. Vietnam Acad. Sci. Technol. **49**(04), 27–40 (2011)
7. Ho, N.C., Lan, V.N., Viet, L.X.: Optimal hedge-algebras-based controller: design and application. Fuzzy Sets Syst. **159**(8), 968–989 (2008)
8. Yager, R.R.: Aggregation operators and fuzzy system modelling. Fuzzy Sets Syst. **67**, 129–145 (1994)

k-Nearest Neighbour Using Ensemble Clustering Based on Feature Selection Approach to Learning Relational Data

Rayner Alfred[1(✉)], Kung Ke Shin[1], Mohd Shamrie Sainin[2], Chin Kim On[1], Paulraj Murugesa Pandiyan[3], and Ag Asri Ag Ibrahim[1]

[1] Faculty of Computing and Informatics,
Universiti Malaysia Sabah, Kota Kinabalu, Sabah, Malaysia
{ralfred,chinkimon,awgasri}@ums.edu.my, keshincode9@gmail.com
[2] School of Computing, Universiti Utara Malaysia, Changlun, Kedah, Malaysia
shamrie@uum.edu.my
[3] School of Mechatronic Engineering,
Universiti Malaysia Perlis, Arau, Perlis, Malaysia
paul@unimap.edu.my

Abstract. Due to the growing amount of data generated and stored in relational databases, relational learning has attracted the interest of researchers in recent years. Many approaches have been developed in order to learn relational data. One of the approaches used to learn relational data is Dynamic Aggregation of Relational Attributes (DARA). The DARA algorithm is designed to summarize relational data with one-to-many relations. However, DARA suffers a major drawback when the cardinalities of attributes are very high because the size of the vector space representation depends on the number of unique values that exist for all attributes in the dataset. A feature selection process can be introduced to overcome this problem. These selected features can be further optimized to achieve a good classification result. Several clustering runs can be performed for different values of k to yield an ensemble of clustering results. This paper proposes a two-layered genetic algorithm-based feature selection in order to improve the classification performance of learning relational database using a k-NN ensemble classifier. The proposed method involves the task of omitting less relevant features but retaining the diversity of the classifiers so as to improve the performance of the k-NN ensemble. The result shows that the proposed k-NN ensemble is able to improve the performance of traditional k-NN classifiers.

Keywords: Relational data mining · k-Nearest Neighbours · Classification · Ensembles · Feature selection · Genetic Algorithm

1 Introduction

Relational data mining is different from traditional data mining methods [1], in which all features obtained from multiple tables that exist in structured

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_35

relational database can be collected, selected and exploited during the data mining process. In other words, a relational data mining involves learning a target table. This target table has a one-to-many relationship with records stored in non-target tables in which patterns across multiple tables can be learnt and thus learning a relational data has the potential to outperform traditional data mining techniques in many cases. This might seem to be an ideal approach to discover useful and interesting patterns from relational databases. Unfortunately when it comes to learning big relational databases with a high degree of one-to-many association, joining features from multiple tables may cause information loss. Therefore, data transformation becomes a tedious trial-and-error work and the classification result is often not very promising especially when the number of tables and the degree of one-to-many association are large. An ensemble of classifiers is a collection of multiple classifiers which is a powerful technique used to improve overall predictive accuracy by consolidating various diversities and accuracies between the classifiers [2,4–6]. Diversified classifiers that do not make the same errors are shown having the abilities to improve performance of ensembles [3]. Sampling from the original datasets and training the classifiers with these datasets obtained from it is the most straightforward to have the classifiers made uncorrelated errors.

Improving the performance of ensemble systems has been attractive because when it comes to classification, ensembles perform better than a single classifier in many cases. A feature selection approach for ensembles has also become very crucial. This is because using good or effective feature selection techniques is able to produce good classification results by selecting only part of the features to be included in the classification process [8]. Selecting different features for different classifiers may provide diversities to the classifiers used for the ensemble system which is a very important aspect to ensemble as aforementioned. There are two approaches of feature selection commonly used which are the filter approach and the wrapper approach [14]. Filter approach selects features by evaluating them with criterion without taking in the predictor used into consideration. In the wrapper approach, the attributes selection process is conducted based on the performance of the predictor which means every subset of features selected will be run on the model for the performance measurement. In this work, a Genetic Algorithm (GA) [8,9] based wrapper feature selection is implemented in order to select the best feature subsets for the k-NN classifiers. In this work, a two-layered GA-based feature selection method is proposed that takes the diversity of the classifiers into consideration in order to omit features that do not carry useful information in the classification task for most of the classifiers. Such approach in reduction of the dimensionality is able to discover important features for classification and thus improving the performance of ensembles.

The rest of this paper is organized as followed. Section 2 explains some related works. Section 3 discusses the proposed two-layered GA based k-NN ensemble classifier based on feature selections order to improve the predictive accuracy of the traditional k-NN classifiers. Section 4 discusses the experimental setup and discusses the experimental results obtained in investigating the effects of performing logical operations to the features selected in the training datasets. Section 5 concludes this paper.

2 Related Works

In relational data mining, the databases involved consist of a collection of data stored in a set of table. Many approaches have been applied in learning relational databases and some of these approaches are applied with ensemble by researchers to improve the performance of learning relational data. Probabilistic Relational Models (PRMs) is a method extended from Bayesian networks and designed for relational learning that looks for good dependency structures that defines the relations between variables in tables from training databases in order to handle relational data [11,16]. A Relational Neighbor (RN) classifier [11] is a simple method that adopts the idea of "guilty by association" which makes predictions on relational data only based on class labels of related neighbors. RN is able to perform competitively when compared to other relational classifiers including PRM. Random forest [18] is applied as classifiers within a hybrid relational learning framework which use both local attributes and flattened (aggregated) relational attributes. Their studies have shown that the prediction accuracy of the ensemble is usually better than individual classification tree.

The motivation of our work is based on previous works related to data summarization [16], conducted by Alfred. In his work, a method called Dynamic Aggregation for Relational Attributes (DARA) is proposed in order to summarize data stored in relational databases that consist of data with one-to-many relations [15,16,22] In DARA algorithm, the entire contents of non-target tables are summarized with respect to the data stored in the target table. The relational data representation is transformed into a vector space representation. Then each feature extracted from the relational model will go through model conversion and computation of component magnitude. Then, the data summarization will be performed in which records stored in non-target table are clustered and will be given a label to indicate the group that the records belong to. Finally, this cluster label is then appended to the target table as an additional column or as a new feature. The empirical results obtained show that DARA algorithm is able to improve the predictive accuracies of C4.5 classifier compared to other relational data mining methods. However, DARA has a major drawback in which the vector space dimensionality will grow larger because it is affected by the large number of distinct values that exist in each column of the relational database. As a result, a feature transformation has been proposed [13,25]. This work is an extended work from [26], in which the discretization of continues values is performed on the record-pattern matrix that consists of the Term Frequency-Inverse Document Frequency (TF-IDF) values. The TF-IDF is used as a statistical measure which expresses the importance of a feature with respect to the class labels. Therefore, the dimensionality of the record-pattern matrix can be reduced because all numerical values are required to be discretized before the feature selection process can be performed based on the feature scoring. In this work, a classification task is performed using ensemble of k-Nearest Neighbors classifiers after summarizing data using DARA with the aforementioned feature transformation process.

3 Ensemble k-Nearest Neighbours Classifier

The k-Nearest Neighbors (k-NN) is a robust and one of the simplest classification methods used to perform a non-parametric classification. The k-NN classifier adopts the concept of classifying a new entry into the same class labels by referring to its k nearest neighbors (entries) in term of their similarity. The Euclidean distance method is commonly chosen to be the proximity measurement among instances where the Euclidean distance, d, between two points p and q in n-dimensional space can be presented as follows,

$$d(p,q) = \sqrt{(q_1 - P_1)^2 + (q_2 - P_2)^2 + \cdots + (q_n - P_n)^2} \tag{1}$$

Despite improving the performance of various classifiers, bagging and boosting ensembles methods are found to be ineffective in improving the performance of k-NN classifiers [4]. Perturbing the training set by bootstrap sampling give not much variation to the nearest neighbor classifier. Thus using these methods on k-NN classifiers are not effective. Ensemble algorithms can be implemented by varying the input selected. Each k-NN classifier using only random subset of features is able to increase the diversity of the classifiers and there are some improvements on the predictive accuracies obtained in various dataset [10]. Despite using a random subset of features that is able to increase the diversity for the k-NN classifiers, it can cause error rate increases significantly when some highly related features are not selected [10]. Such condition will result in producing poor classifiers and thus making the ensembles inefficient. A genetic algorithm (GA) can be used to optimize the feature selection process [12–14]. Thus, the aim of this work is to investigate the performance of the k-NN classifiers when the consensus features of the k-NN classifiers are obtained by applying logical operations. The feature selection process can be divided into two stages. In stage 1, a genetic algorithm [8] is applied to search for the best set of features for each value of k used for the k-NN classifiers and all these set of features are used as the input for stage 2. A two third validation technique is applied in conducting the experiments. The fitness function of the GA based feature selection process is the predictive accuracy of the k-NN classifier used. Figure 1 illustrates the process of selecting features by using the k-NN ensemble classifier in order to obtain consensus features from multiple classifiers. Given m features that exist in the dataset, Fig. 2 illustrates a set of selected features are collected for each value k of the k-NN classifier used to form the input for the second stage of the feature selection process.

In stage 2, c number of k-NN classifiers are selected in producing the final set of features, where $c \leq k$ and k refers to the k-NN classifiers (e.g., k is 5 when 5-NN classifier is used. A genetic algorithm is used in order to select the best set of c number of k-NN classifiers before the final set of features can be produced. The chromosome used in stage 2 is different compared to the chromosome used in stage 1. In stage 1, the size of the chromosome is m (number distinct features in the dataset), whereas in stage 2, the size of the chromosome is $k+2$ (see Fig. 3), where k is the number of k-NN classifiers used in stage 1 and the last 2

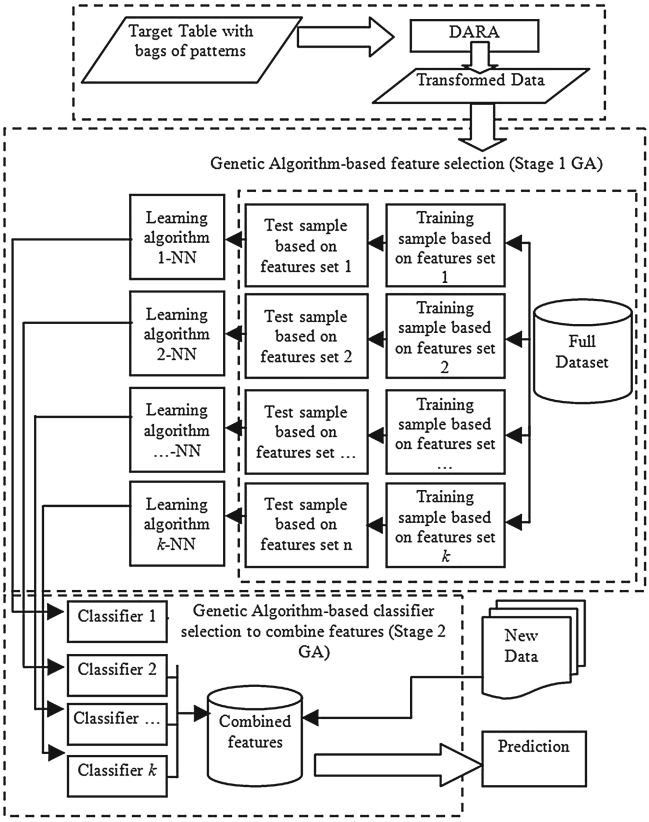


Fig. 1. Two-layered Genetic Algorithm.

k-NN	F_1	F_2	F_3	F_4	F_5	…	F_m
1	1	0	0	0	0	…	1
2	0	0	1	1	0	…	1
…	…	…	…	…	…	…	…
k	1	0	0	0	0	…	0

Fig. 2. Features selected for each k of k-NN classifier.

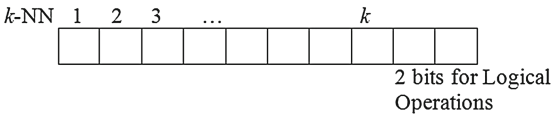


Fig. 3. Chromosome used in stage 2 GA.

genes of the chromosome refers to the type of logical operations that is used to find consensus decision on the final set of features selected. The results obtained in stage 1 will be used in stage 2, where each chromosome is represented as a string of k + 2 bits of 0 and 1, where k is the number of k-NN classifiers. In stage 2, there are c unique sets of features that will be selected by referring to the chromosome used and the final set of features selected will be further determined by performing the logical operation to obtain a new set of features.

4 Experimental Design and Results

4.1 Experimental Setup

The experiments conducted in this study can be divided into two (2) separate phases in this study. The main task in the first part of the study is to select the best subsets of features (from m distinct features) for every k-NN classifier where $k = n^{\frac{1}{2}}$ and n is the number of instances that exist in the dataset. The chromosome's size used in stage 1 equals to the total number of features, m, in the dataset and the chromosome's size used in stage 2 equals to the number of classifier, k + 2, in which there are two additional bits added. A single point crossover is used for the crossover process and the other parameters include the population's size of 50, the number of generations of 50, and mutation's rate of 2 %. A two third validation will be used to evaluate the predictive accuracy for the given Mutagenesis data and Hepatitis data. Figure 1 illustrates the processes involved starting from summarizing data to extracting the final consensus set of features from classifiers in this experiment (Fig. 4).

In stage 2, the chromosome represents a set of k-NN classifiers that are selected and coupled with the sets of selected features obtained previously. All features of the selected classifiers will undergo one of the three different types of logical operations which are AND, OR, and XOR operations. The type of the logical operation that will be performed is determined by the values of the last

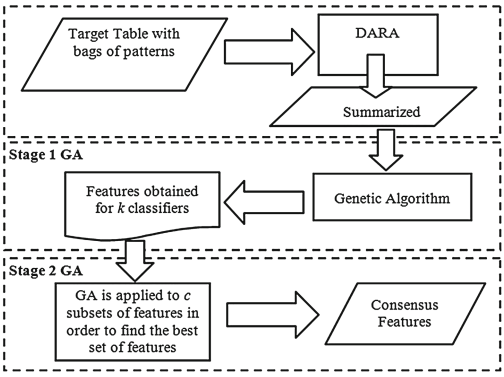


Fig. 4. Two-layered GA is applied to obtain consensus.

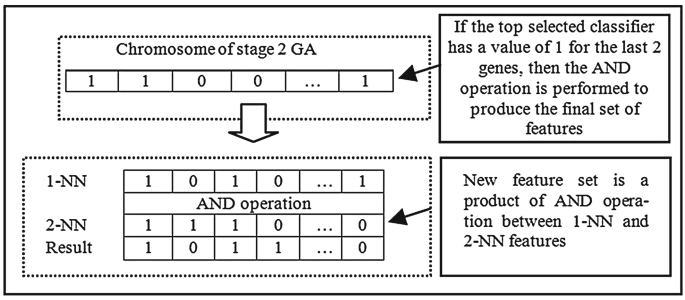


Fig. 5. A set of features selected is produced by performing the logical operation in stage 2 GA.

two genes of the chromosome. When the last two genes are both '0', then the OR operation will be performed. The AND operation will be performed if the last two genes are both "1", whereas the XOR operation will be performed when the values of the genes are '0' and '1' or vice-versa. For instance, there will be k different sets features selected for each k-NN classifier used in stage 1. The set of features selected for each value of k is optimized using a GA based on the prediction accuracy of the k-NN classifier. Based on the results obtained in stage 1, a set of c number of k-NN classifiers is then selected in order to produce the final set of selected features, where $c \leq k$. The features of selected classifiers, for example features obtained from the 1-NN and 2-NN, will be taken and the logical operation will be performed in order to generate or produce a new set of features as shown in Fig. 2, where the type of logical operation performed is based on the last two genes of the chromosome. The percentage of selecting the subsets is kept particularly low because too many subsets selected will make the logical operation meaningless since it will cause most of the outputs obtained by performing the AND operation to be false and most of the outputs obtained by performing the OR operation to be true. Table 1 outlines the predictive accuracy results of the k-Nearest Neighbor classifier on the three mutagenesis datasets (B1, B2, and B3) and hepatitis datasets (H1, H2, and H3) (Fig. 5).

4.2 Mutagenesis and Hepatitis Datasets

Two sample databases (mutagenesis and hepatitis databases) are described in this subsection.

The data in mutagenesis domain [27] describes 188 molecules falling in two classes, mutagenic (active) and non-mutagenic (inactive) and 125 of these molecules are mutagenic. The description consists of the atoms and bonds that make up the compound. Thus, a molecule is described by listing its atoms *atom(AtomID, Element, Type, Charge)* and the bonds *bond(Atom1, Atom2, BondType)* between atoms. In this experiment, three different sets of background knowledge (B1, B2 and B3) will be used.

- B1: The atoms in the molecule are given, as well as the bonds between them; the type of each bond is given as well as the element and type of each atom.
- B2: Data in B1 and continuous values about the charge of atoms are added
- B3: Data in B2 and two continuous values describing each molecule are added, which are the log of compound's octanol/water partition coefficient ($logP$) and en-ergy of the compounds lowest unoccupied molecular orbital ($^{\varepsilon}LUMO$).

The second relational dataset is called hepatitis databases, which is obtained from PKDD 2005. Three different sets of background knowledge will be used in all the experiments performed in this work. They will be referred to as datasets H1, H2 and H3:

- H1: The hematological examinations are given, which include the examinations for white blood cells count (WBC), red blood cells count (RBC), hemoglobin count (HGB) and hematocrit (HCT). White blood cells (WBC) or leukocytes are cells of the immune system defending the body against both infectious disease and foreign materials.
- H2: In addition to the data in dataset H1, continuous values describing the GOT, GPT, LAP, IBIL, CRE examinations from the in-hospital examinations are added.
- H3: In addition to the data in dataset H2, continuous values describing the TSH, SEC examinations from the out-hospital examinations are added.

4.3 Results

The result shows that the average predictive accuracies of the classification task obtained are improved though not significant. Based on the results shown in Table 1, it can be shown that the proposed 2-layered Genetic Algorithm based framework is able to boost the predictive accuracy. This is due to the fact that, most relevant features are selected based on its frequencies of being selected for every value of k when the k-NN classifier is executed. As a result, all features which are more persistence to be selected in most cases will have higher chances to be selected in the second layer of the GA based framework.

Table 1. Average classification accuracies for k-NN

Dataset	With GA only	With proposed 2-layered GA
B1	87.4	88.8
B2	88.4	88.7
B3	92.1	92.4
H1	69.9	70.0
H2	66.6	67.1
H3	69.5	69.7

5 Conclusion

This paper investigates the effect of applying ensemble of k-NN classifiers to the data summarized using DARA algorithm in order to learn relational data. Our main objective is to extract the best subsets of features by obtaining consensus features from ensemble of classifiers. Results obtained from the experiments conducted show that by using the proposed method to obtain consensus features from different classifiers can increase the classification performances. However, the improvements are not significant. This could be due to the lack of diversities measurements in the first stage of GA. In future work, more investigations will be required to tune different parameters such as fitness, measurement methods of similarities between instances and also the measurement of diversities between classifiers.

Acknowledgments. This work has been supported by the Research Grant Scheme project funded by the Ministry of Education (MOE), Malaysia, under Grants No. FRG0382-ICT-2/2014.

References

1. Fayyad, U., Shapiro, G.P., Smyth, P.: From data mining to knowledge discovery in data mining. AI Mag. **17**(3), 37–54 (1996)
2. Hansen, L.K., Salamon, P.: Neural network ensembles. IEEE Trans. Pattern Anal. Mach. Intell. **12**, 993–1001 (1990)
3. Ali, K.M., Pazzani, M.J.: Error reduction through learning multiple descriptions. Mach. Learn. **24**, 173–202 (1996)
4. Breiman, L.: Bagging predictors. Mach. Learn. **24**, 123–140 (1996)
5. Freund, Y., Schapire, R.: Experiments with a new boosting algorithm. In: Thirteenth International Conference on Machine Learning (1996)
6. Quinlan, J.R.: Bagging, boosting and C4.5. In: Fourteenth National Conference on Artificial Intelligence (1996)
7. Cover, T.M., Hart, P.E.: Nearest neighbor pattern classification. IEEE Trans. Inf. Theory **13**(1), 21–27 (1967)
8. Holland, J.: Adaptation in Natural and Artificial Systems. University of Michigan Press, Ann Arbor (1975). MIT Press, Cambridge (1992)
9. Fraser, A.S.: Simulation of genetic systems by automatic digital computers I. Introduction/Aust. J. Biol. Sci. **10**, 484–491 (1957)
10. Bay, S.D.: Nearest neighbour classification from multiple feature subsets. Intell. Data Anal. **3**(3), 191–209 (1999)
11. Getoor, L.: Multi-relational data mining using probalilistic relational models: research summary. In: Proceedings of the First Workshop in Multi-Relational Data Mining (2001)
12. Xia, P.Y., Ding, X.Q., Jiang, B.N.: A GA-based feature selection and ensemble learning for high-dimensional datasets. IEEE Int. Conf. Mach. Learn. Cybern. **3**, 7–12 (2009)
13. Canuto, A.M.P., Nascimento, D.S.C.: A genetic-based approach to features selection for ensembles using a hybrid and adaptive fitness function. In: IEEE International Joint Conference on Neural Networks (IJCNN), pp. 1–8 (2012)

14. Guyon, I., Elisseeff, A.: An introduction to variable and feature selection. JMLR **3**, 1157–1182 (2003)
15. Saeys, V., Inza, I., Larrañaga, P.: A review of feature selection techniques in bioinformatics. Bioinformatics **23**(19), 2507–2517 (2007)
16. Ghanem, A.S., Venkatesh, S., West, G.: Learning in imbalanced relational data. In: 2008 19th International Conference on Pattern Recognition, ICPR 2008, pp. 1–4 (2008)
17. Macskassy, S., Provost, F.: A simple relational classifier. In: Proceedings of 2nd Workshop on Multi-Relational Data Mining (MRDM) (2003)
18. Chen, J.X., Li, P.B.: Random forest for relational classification with application to terrorist profiling. In: IEEE International Conference on Granular Computing, GRC 2009, pp. 630–633 (2009)
19. Alfred, R.: Optomizing feature construction process for dynamic aggregation of relational attributes. J. Comput. Sci. **5**(11), 864 (2009)
20. Kheau, C.S., Alfred, R., Keng, L.H.: Dimensionality reduction in data summarization approach to learning relational data. In: Selamat, A., Nguyen, N.T., Haron, H. (eds.) ACIIDS 2013. LNCS (LNAI), vol. 7802, pp. 166–175. Springer, Heidelberg (2013). doi:10.1007/978-3-642-36546-1_18
21. Kuncheva, L., Jain, L.: Designing classifier fusion systems by genetic algorithms. IEEE Trans. Evol. Comput. **4**(4), 327–336 (2000)
22. Alfred, R.: The study of dynamic aggregation of relational attributes on relational data mining. In: Alhajj, R., Gao, H., Li, J., Li, X., Zaïane, O.R. (eds.) ADMA 2007. LNCS (LNAI), vol. 4632, pp. 214–226. Springer, Heidelberg (2007). doi:10.1007/978-3-540-73871-8_21
23. Alfred, R., Kazakov, D.: A clustering approach to generalized pattern identification based on multi-instanced objects withDARA. In: Local Proceedings of ADBIS, Varna, pp. 38–49 (2007)
24. Alfred, R., Kazakov, D.: Pattern-based transformation approach to relational domain learning using DARA. In: Crone, S.F., Lessmann, S., Stahlbock, R. (eds.) The Proceedings of the 2006 International Conference on Data Mining (DMIN 2006), 25–29 June, pp. 296–302. CSREA Press, Las Vegas (2006). ISBN: 1-60132-004-3
25. Alfred, R.: Feature transformation: a genetic-based feature construction method for data summarization. Comput. Intell. **26**(3), 337–357 (2010)
26. Alfred, R., Kazakov, D.: Discretization numbers for multiple-instances problem in relational database. In: Ioannidis, Y., Novikov, B., Rachev, B. (eds.) ADBIS 2007. LNCS, vol. 4690, pp. 55–65. Springer, Heidelberg (2007). doi:10.1007/978-3-540-75185-4_6
27. Srinivasan, A., Muggleton, S.H., Sternberg, M.J.E., King, R.D.: Theories for mutagenicity: a study in first-order and feature-based induction. Artif. Intell. **85**, 277–299 (1996)

Low Power ECC Implementation on ASIC

Van-Lan Dao(✉), Van-Tinh Nguyen, and Van-Phuc Hoang

Le Quy Don Technical University, Hanoi, Vietnam
kqha1025@gmail.com, tinh.vxl@gmail.com, phuchv@mta.edu.vn

Abstract. In this paper, the Low power Elliptic Curve Cryptography (ECC) structure over Galois field $GF(2^m)$ is studied and implemented on the Application Specific Integrated Circuit (ASIC) tool for wireless sensor network and Internet of Things (IoT) Applications. Clock gating technique is used for decreasing power consumption. The implementation is conducted by the 180 nm CMOS standard library shows that the proposed ECC structure has the power consumption of 10.4 μW/MHz outweigh than previous designs.

1 Introduction

Recently, the need of using energy-efficient electronic devices, IoT applications and communication networks is more and more emerging. A large number of applications using wireless sensor networks to be used in different areas of life. Wireless sensor networks [2–7] can be utilized in military, agriculture, industry, medicine, sport, environmental monitoring, traffic, smart houses, etc. Figure 1 is the general structure of a wireless sensor network (WSNs) in which the data confidentiality is an essential issue [1].

The objective of this paper is to design a low power ECC structure based on the clock gating technique with area and power constrained condition. The rest of this paper is organized as follows. Section 2 describes the ECC, clock gating technique and Sect. 3 presents the low power ECC structure. Section 4 shows the implementation results and finally, Sect. 5 concludes the paper.

2 ECC and Clock Gating Technique

2.1 Elliptic Curve Cryptography

In 1985, elliptic curve cryptography [15] based on the discrete logarithm problem was proposed by Miller [8] and Koblitz [9] independently. An elliptic curve, over binary field ($GF(2^m)$) can be defined a set of solution to the equation as bellow [10]:

$$y^2 + xy = x^3 + ax^2 + b \tag{1}$$

where $a, b \in GF(2^m), b \neq 0$ and at the same time the point at infinity is $\emptyset$. Thus, if P_1 is a point on the ECC curve then $P_1 + \emptyset = P_1$. Two point operations are used called point addition and point doubling. We consider P_1 and P_2 be points

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_36

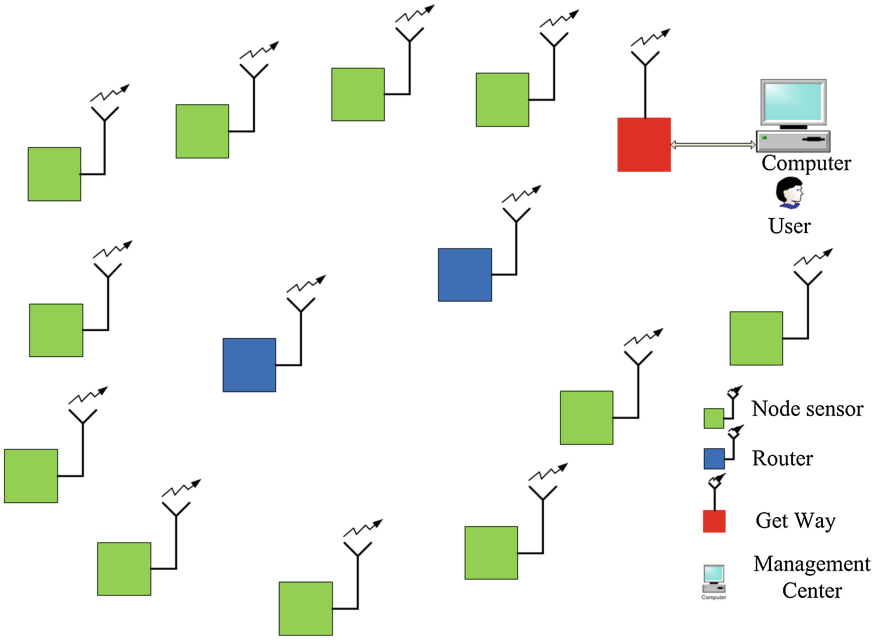


Fig. 1. The general structure of a wireless sensor network.

over $(GF(2^m)$ on the elliptic curve. If the two points are not the same then the adding of the points gives a new point, Q on the elliptic curve called the result of point addition. Again, if the two points are the same, $P_1 = P_2$ then, the addition of the two points gives a new point, Q is called the result of point doubling [10].

The main operation in ECC is point multiplication, where a point, P and a random number, k are considered on an ECC curve and a scalar multiplication is performed to obtain Q, where $Q = k.P$. The basic operation of point multiplication can be obtained by point addition and point doubling. The field arithmetic involved in point addition and point doubling can also be performed in binary fields. The scalar point multiplication $Q = k.P$ can be written as $kP = P + \cdots + P$.

2.2 Clock Gating Technique

Today there are many researches related to clock gating technique [11–13]. Clock gating technique is an efficient power optimization technique that is employed in both ASIC and FPGA designs to eliminate the unnecessary switching activity. This method usually requires the designers to add a small amount of logic to their RTL code to deselect unnecessarily active sequential elements - registers. Figure 2 is the simple classic global clock gating.

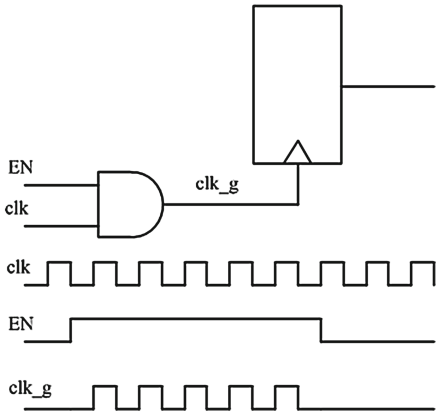


Fig. 2. The simple global clock gating.

3 Low Power ECC Design

In our research, the ECC core architecture as shown in Fig. 3 is used [14]. It includes a FSM control unit, point addition units, point multiplication units and squaring units.

Table 1 lists the function of the signals in the proposed ECC core. On the other hand, clock gating technique is carried out by using the Multiplexer unit.

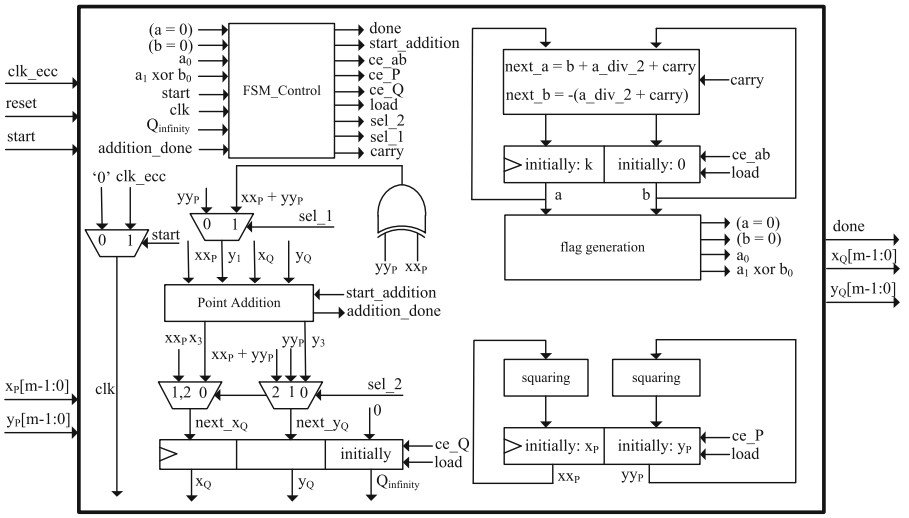


Fig. 3. The ECC core architecture.

Table 1. Signals in the proposed Ecc core.

Signal	Direction	Function
clk_ecc	Input	System clock
reset	Input	System reset
$X_P[m-1:0]$	Input	Data input X_P
$Y_P[m-1:0]$	Input	Data input Y_P
done	Output	To indicate that the output is ready to read
$X_Q[m-1:0]$	Output	Data output X_Q
$Y_Q[m-1:0]$	Output	Data output Y_Q

4 Implementation Results

The ECC core was implemented with VHDL code, simulated in the Modelsim tool and then synthesized by using a 180 nm CMOS standard library in the Synopsys Design Complier. Figure 4 shows the simulation model for the 8-bit ECC core. The input generation block generates the input vector values for the ECC core verification. Figures 5 and 6 present the RTL (pre-synthesis) simulation results in the Modelsim tool and post-synthesis simulation results in the Synopsys VCS-DVE tool, respectively. Table 2 is an example of a test vector for the ECC core verification. It can be seen that the simulation results are the same for pre- and post-synthesis netlists.

Table 2. An example of a test vector for ECC core verification.

K(hexa)	0xFFF030001F0000FFFFF000003800000000
X_P**(hexa)**	0x2FE13C0537BBC11ACAA07D793DE4E6D5E5C94EEE8
Y_P**(hexa)**	0x289070FB05D38FF58321F2E800536D538CCDAA3D9
X_Q**(hexa)**	0x01C14DDAB12BC0D98BF83CE0022F305039F64FC205
Y_Q**(hexa)**	0x006F9B20200EB3CEA80D1DB6C0FB8E6DED4A3C665C

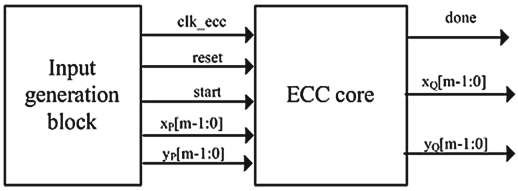


Fig. 4. The simulation model for the ECC core.



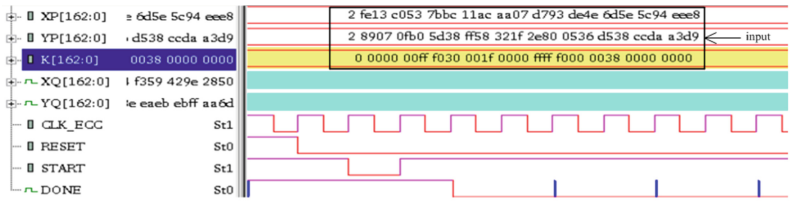
(a) Data and key input



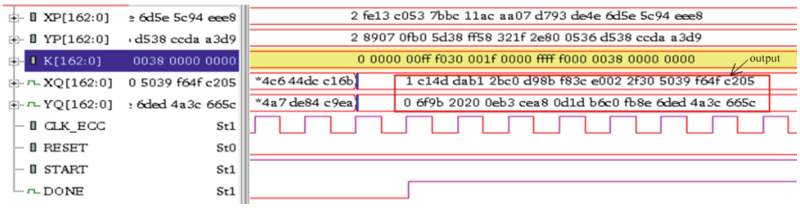
(b) Data output.

Fig. 5. Simulation results in Modelsim tool

The ASIC implementation results are shown in Table 5 in which the proposed ECC core power can be reduced to only 10.4 µW/MHz when $m = 163$. Compared with other designs, the ECC core power consumption can also be reduced significantly while achieving the maximum clock frequency of 59 MHz.



(a) Data and key input



(b) Data output.

Fig. 6. Post-synthesis simulation results in Synopsys VCS-DVE tool

Table 3. Implementation results of ECC core in a 180 nm CMOS ASIC library.

	K163	K233	K283	K409	K571
Area (kgates)	27.5	37.2	47.3	65.5	98
Speed (MHz)	59	56	61	56	55

Table 4. Power consumption of ECC core in a 180 nm CMOS ASIC library (μW).

ECC core	K163	K233	K283	K409	K571
Not clock gating	143	203	246	356	499
Clock gating	10.4	40	58	108	174

Table 5. Comparison of ECC core designs.

Design	Field size (m)	Power	Note
[16]	163	17 µW	4-bit digit multiplier
		305 µW	3-reg. coprocessor
		503 µW	7-reg. coprocessor
		12000 µW	typ. 8-bit processor
[17]	233	20.38 mW	Single Processor
		64.64 mW	Asynchronous MIMD
		49.51 mW	MIMD-SIMD
[18]	192	39.3 µW	-
[19]	192	18.85 µA@100 kHz	-
[20]	134	<15 µW@200 kHz	-
[21]	100	<400 µW@500 kHz	-
[22]	193	42.8 µW/MHz	-
This work	233	40 µW/MHz	-

In this paper, the ECC core is synthesized with two cases of using clock gating techniques and not using clock gating techniques. Area and speed of the ECC core in both cases were similar, synthesis results are shown in Table 3. However, the power consumption of the ECC core with clock gating technique is reduced significantly. For example, $m = 163$, the power consumption of ECC core using clock gating technique decreased by 92.7 % in comparison to the case not using clock gating technique. Other cases of m are shown in Table 4, in which K163 presents the key length of 163-bit. Figure 7 is the synthesized netlist in the 180 nm CMOS standard cell library for the case of clock gating architecture.



Fig. 7. Synthesized netlist of the ECC core in a 180 nm CMOS library.

5 Conclusions

This paper presented a power efficient ECC core for wireless networks and IoT applications. The ASIC implementation results show that using clock gating techniques, the ECC core power can be reduced significantly. Therefore, this ECC core is highly potential to be used in IoT applications and wireless network nodes such as environment monitoring which requires low power and compact encryption cores. In the future, we will optimize the power consumption and area for the proposed ECC core and apply it for wireless network applications.

References

1. Du, X., Chen, H.-H.: Security in wireless sensor networks. IEEE Wireless Commun. **15**(4), 60–66 (2008)
2. Bari, N., Mani, G., Berkovich, S.: Internet of things as a methodological concept. In: 2013 Fourth International Conference on Computing for Geospatial Research and Application (COM.Geo), San Jose, CA, pp. 48–55 (2013)
3. Hac, A.: Wireless Sensor Network Designs. Wiley, New York (2013)
4. Tiwari, A., et al.: Energy-efficient wireless sensor network design and implementation for condition-based maintenance. ACM Trans. Sensor Netw. (TOSN) **3**(1), 1–23 (2007)
5. Yun, D.S., Lee, S.-J., Kim, D.H.: A study on the vehicular wireless base-station for in-vehicle wireless sensor network system. In: Proceedings of 2014 International Conference on Information and Communication Technology Convergence, pp. 609–610, October 2014

6. Ding, R., Hou, J., Xing, B.: Research of wireless sensor network nodes based on ambient energy harvesting. In: Proceedings of 2013 6th International Conference on Intelligent Networks and Intelligent Systems (ICINIS), pp. 286–288, November 2013
7. Liu, Y., Wang, C., Qiao, X., Zhang, Y., Yu, C.: An improved design of ZigBee wireless sensor network. In: Proceedings of 2nd IEEE International Conference on Computer Science and Information Technology (ICCSIT 2009), pp. 515–518, August 2009
8. Miller, V.S.: Use of elliptic curves in cryptography. In: Williams, H.C. (ed.) CRYPTO 1985. LNCS, vol. 218, pp. 417–426. Springer, Heidelberg (1986). doi:10.1007/3-540-39799-X_31
9. Koblitz, N., Menezes, A., Vanstone, S.: The state of elliptic curve cryptography. Des. Codes Crypt. **19**(2–3), 173–193 (2000)
10. Hankerson, D., Menezes, A., Vanstone, S.: Guide to Elliptic Curve Cryptography. Springer, New York (2004)
11. Zhong, W.J., Noh, N.M., Rosdi, B.A.: Clock gating assertion check: an approach towards achieving faster verification closure on clock gating functionality. In: 2015 6th Asia Symposium on Quality Electronic Design (ASQED), Kula Lumpur, pp. 94–101 (2015)
12. Wimer, S., Koren, I.: Design flow for flip-flop grouping in data-driven clock gating. IEEE Trans. Very Large Scale Integr. (VLSI) Syst. **22**(4), 771–778 (2014)
13. Nikolić, M., Katona, M.: Improve the automatic clock gating insertion in ASIC synthesis process using optimal enable function selection. In: 2010 5th European Conference on Circuits and Systems for Communications (ECCSC), Belgrade, pp. 131–134 (2010)
14. Deschamps, J.-P., Imaãa, J.L., Sutter, G.D.: Hardware Implementation of Finite-Field Arithmetic, 1st edn. Publisher of McGraw-Hill Education, New York (2009)
15. National Institute of Standards and Technology (NIST): Digital Signature Standard (DDS). FIPS Publication 186-3, June 2009
16. Bertoni, G., Breveglieri, L., Venturi, M.: ECC hardware coprocessors for 8-bit systems, power consumption considerations. In: Third International Conference on Information Technology: New Generations, 2006, ITNG 2006, Las Vegas, NV, pp. 573–574 (2006)
17. Purnaprajna, M., Puttmann, C., Porrmann, M.: Power aware reconfigurable multiprocessor for elliptic curve cryptography. In: Design, Automation and Test in Europe, 2008, Munich, pp. 1462–1467 (2008)
18. Ahmadi, H.R., Afzali-Kusha, A.: Low-power low-energy prime-field ECC processor based on montgomery modular inverse algorithm. In: 12th Euromicro Conference on Digital System Design, Architectures, Methods and Tools, 2009, Patras, pp. 817–822 (2009)
19. Feldhofer, M., Wolkerstorfer, J.: Strong crypto for RFID tags - comparison of low-power hardware implementations. In: ISCAS 2007, pp. 1839–1842 (2007)
20. Batina, L., et al.: Public-key cryptography on the top of a needle. In: ISCAS 2007, pp. I831–1834 (2007)
21. Gaubatz, G., et al.: State of the art in ultra-low power public key cryptography for wireless sensor networks. In: 3rd IEEE PerCom 2005 Workshops, pp. 146–150 (2005)
22. de Dormale, G.M., Ambroise, R., Bol, D., Quisquater, J.J., Legat, J.D.: Low-cost elliptic curve digital signature coprocessor for smart cards. In: Proceedings of the IEEE 17th International Conference on Application-Specific Systems, Architectures and Processors, ASAP 06, pp. 347–353 (2006)

Malwares Classification Using Quantum Neural Network

Tu Tran Anh(✉) and The Dung Luong

Academy of Cryptography Techniques,
141, Chien Thang Str., Tan Trieu, Thanh Tri, Hanoi, Vietnam
trananhtu.1407@gmail.com, thedungluong@gmail.com

Abstract. Quantum neural networks (QNNs) have been explored as one of the best approach for improving the computational efficiency of neural networks. Because of the powerful and fantastic performance of quantum computation, some researchers have begun considering the implications of quantum computation on the field of artificial neural networks (ANNs). The purpose of this paper is to introduce an application of QNNs in malwares classification. Inherently Fuzzy Feedforward Neural Networks with sigmoidal hidden units was used to develop quantized representations of sample information provided by the training data set in various graded levels of certainty. Experimental results presented here show that (QNN's) gave a kind of fast and realistic results compared with the (ANN's). Simulation results indicate that QNN is superior (with total accuracy of 98.245 %) than ANN (with total accuracy of 95.214 %).

Keywords: Artificial neural network · Quantum computing · Quantum neural network · Malware classification

1 Introduction

One of the most serious problems facing the Internet today is malicious software. Spam, phishing, denial of service attacks, botnets, and worms largely depend on some form of malicious code, commonly referred to as malware. Malware is often used to infect the computers of unsuspecting victims by exploiting software vulnerabilities or tricking users into running malicious code. Understanding this process and how attackers use the backdoors, key loggers, password stealers, and other malware functions is becoming an increasingly difficult and important problem. Unfortunately, the complexity of modern malware is making this problem more difficult. Malware are obfuscated, packed and many more. It makes malware polymorphism and metamorphism so that very difficult to detect. Making the problem even more challenging is the increase in the number and diversity of Internet malware. Each day there are thousands of malware variants appear. Microsoft receives over 150 thousand new, unknown files each day to be analyzed. Given the enormous volume of unknown files, manual inspection by analysts is impossible. Malware authors use automated methods such as polymorphism,

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_37

where a program generates a unique, new instance of a malware family for each victim, to create new malware. To combat this threat, anti-virus companies must utilize signal processing and machine learning methods to automatically detect new instances of malware. The goal of a fully automated malware classification system is to operate at an extremely low false positive rate while providing a reasonably low false negative rate. When the false positive is higher, it may result in an important system file being deleted from the computer thereby preventing the computer from booting correctly. And the false negative rate provided that the system predicts a file to be malicious, does it belong to a known malware family? Being able to correctly predict the family allows an unknown file to be assigned to the correct expert for manual investigation. Malware classification is an active research area. Malware classification systems are often based on sparse, binary feature sets [11]. There are many method were used in this area, such as machine learning, datamining [10], and neural network [5,8]. Among these techniques neural network is one of the best for malware classification with error rate only 0.42 % [5]. Many researches have studied to improve the efficiency of neural network [2,4]. One of the best approach is applied quantum computing in artificial neural network [1,6,7,12]. Quantum neural networks (QNNs) have been explored as one of the best approach for improving the computational efficiency of neural networks. It is a promising area in quantum computation and quantum information field. Since K. Grover proposed a method that can speed up a range of search applications over unsorted data using Quantum mechanics [9] several models have been proposed in the literature but most of them need clear hardware requirements to implement such models, one of the most exciting emerging technologies is quantum computation, which attempts to overcome limitations of classical computers by employing phenomena unique to quantum-level events, such as nonlocal entanglement and superposition. It is therefore not surprising that many researchers have conjectured that quantum effects in the brain are crucial for explaining psychological phenomena, including consciousness, Jarernsri. L. Mitrpanont and Ananta Srisuphab, presented the approach of the quantum complex-valued backpropagation neural network or QCBPN in [7], the challenge of their research is the expected results from the development of the quantum neural network using complex-valued backpropagation learning algorithm to solve classification problems. In our work, we employ sparse binary features based on file strings, application programming interface (API) tri-grams, and API call plus parameter value. This called dataset, and transform the binary dataset in to internal presentations in qubit form to train in QNN. We use the standard feed-forward neural network topologies trained with backpropagation. This paper is organized as follows: Sect. 2 describes the QNN and its learning algorithms for effective malware classification. We present our malware classification system based on QNN in Sect. 3. The simulation and result is shown in Sect. 4. We offer limitations, future directions and conclude in Sect. 5.

2 QNN's Learning Algorithm

QNN has n_i inputs, one hidden layer of n_h nodes, each one represents a multilevel units and n_o outputs. Output units are sigmoid. Let $xk = [x_{1,k}, x_{2,k}, ..., x_{n,k}]^T, k = 1, 2, ..., m$ be the k-th sample vector of the data set X. Then the input to the j-th hidden unit from the k-th sample vector is:

$$\overline{h_{j,k}} = \sum_{l=0}^{n_i} v_{jl} x_{l,k} \tag{1}$$

$$\widetilde{h_{j,k}} = \frac{1}{n_s} \sum_{r=1}^{n_s} h_{j,k}^r = \frac{1}{n_s} \sum_{r=1}^{n_s} sgm(\beta_h(\overline{h_{j,k}} - \theta_j^r)) \tag{2}$$

where $\widetilde{h_{j,k}}$ is the output of the j-th multilevel hidden unit from k-th sample, $\text{sgm}(\tau) = \frac{1}{1+e^{-\tau}}$ is a sigmoid function, β_h is a slope factor for all multilevel hidden units, θ_j^r defines a jump positions in transfer function, and n_s is the number of levels or sigmoid in multilevel hidden unit, Fig. 1 show the response $\widetilde{h_{j,k}}$ of j-th four levels quantum neuron as a function of its input $\overline{h_{j,k}}$ with equal step heights [3].

$$\overline{y_{i,k}} = \sum_{j=0}^{n_h} w_{ij} \widetilde{h_{j,k}} \tag{3}$$

We have $\widetilde{h_{0,k}} = 1, \forall k$ therefore, the response of the i-th unit for k-th input sample can be written as follow:

$$\hat{y}_{i,k} = \text{sgm}(\beta_0(\overline{y_{i,k}})) \tag{4}$$

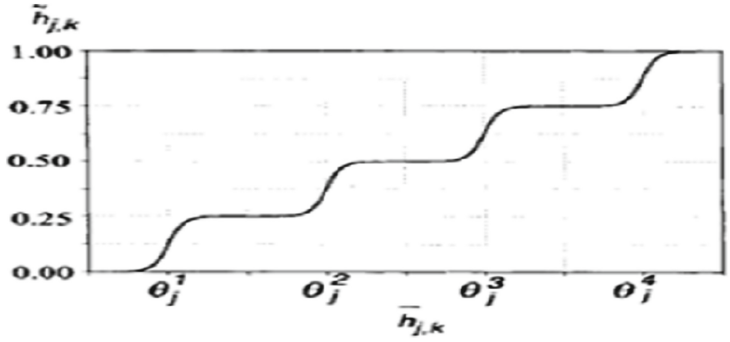


Fig. 1. Multilevel transfer function with equal step heights

The QNN learning algorithm are presented as follows:
Update the quantum intervals:

Step 1: Select $\alpha, \alpha_\theta, \beta_h, \beta_0$ and randomly initialize the weights $W\&V$ and values of jump positions θ_j^r.
Step 2: Present k-th input pattern and specify the desired output
Step 3: Calculate actual output $\hat{y}_{i,k}$, using the present values ofv_{jl} and w_ij.
Step 4: Find the error output $e_{i,k}^o$ and hidden layer error $e_{j,k}^h$.

$$e_{i,k}^o = (y_{i,k} - \hat{y}_{i,k})\hat{y}_{i,k}(1 - \hat{y}_{i,k}) \tag{5}$$

$$e_{j,k}^h = (\frac{1}{n_s}\sum_{r=1}^{n_s} h_{j,k}^r(1 - h_{j,k}^r))\sum_{p=1}^{n_o} e_{p,k}^o w_{pj} \tag{6}$$

p is over all nodes in the layer above node j
Step 5: Adjust the synaptic weights:

$$w_{ij,k} = w_{ij,k-1} + \alpha e_{i,k}^o \frac{1}{n_s}\sum_{r=1}^{n_s} sgm(\beta_h(\overline{h_{j,k}} - \theta_j^r) \tag{7}$$

$$v_{jl,k} = w_{jl,k-1} + \alpha\beta_h(\frac{1}{n_s}\sum_{r=1}^{n_s} h_{j,k}^r(1 - h_{j,k}^r)x_{l,k}) \tag{8}$$

Step 6: Present another input pattern and go back to step 2.

Update the quantum intervals
Step 1: In each training cycle, calculate the outputs $\hat{y}_{i,k}$ and $v_{q,k}^r$ for each hidden node. Then, take the average values for each class, $\langle\tilde{h}_{i,c_m}\rangle$, $\langle v_{i,c_m}\rangle^s$ for m-th class during the training of QNN as:

$$\langle\tilde{h}_{i,c_m}\rangle = \frac{1}{|c_m|}\sum_{x_k \in c_m} \tilde{h}_{j,k} \tag{9}$$

$$\langle v_{i,c_m}\rangle^s \frac{1}{|c_m|}\sum_{x_k \in c_m} v_{j,k}^r \tag{10}$$

Step 2: Calculate the quantum interval adjustment for each level:

$$\Delta\theta_j^r = \alpha_\theta \frac{\beta_h}{\alpha_h}\sum_{m=1}^{n_o}\sum_{x_k \in c_m}\left(\langle\tilde{h_{j,c_m}}\rangle - \tilde{h_{j,k}}\right)\left(\langle v_{j,c_m}^s\rangle - v_{j,k}^s\right) \tag{11}$$

Step 3: Update jump-positions by:

$$\theta_j^r = \theta_j^r + \Delta\theta_j^r \tag{12}$$

Step 4: Continue next cycle and go back to step 2.

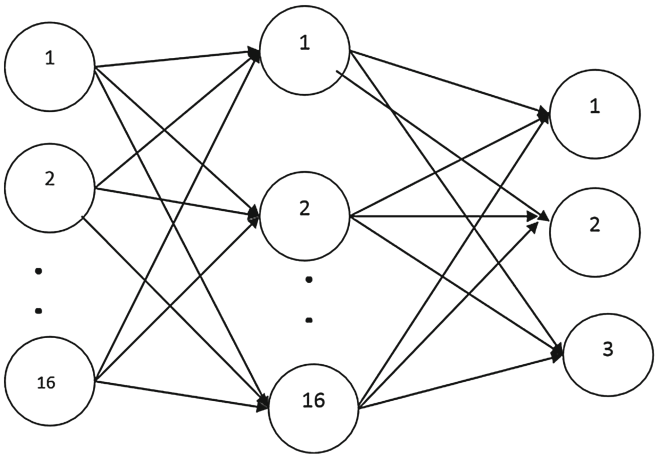


Fig. 2. QNN training system

3 Malware Classification Quantum Neural System

The quantum system we proposed here is a feed-forward network trained with back-propagation. The QNN has 16 input nodes and a hidden layer with 16 nodes, 3 nodes output layer as shown in Fig. 2.

Our dataset is constructed by 900 sample on vxheanven.org includes: 300 viruses, 300 trojans and 300 benign samples. We use 250 samples of each type to construct training set and 50 sample of each for testing set. Each sample is normalized from its API call using tri-gram and API parameter features. After extracting, all sample are transfer in to quantum form and put in QNN.

4 Simulation and Experimental Result

Three layers QNN and three layers ANN was employed as classifier of malware sample, the performance of the QNN was tested to perform classification on malware sample. With 750 sample (250 viruses, 250 trojan, 250 benign samples) and 150 for testing (testing input feature vectors 50 viruses, 50 trojan, 50 benign samples), the structure is 16-16-3 for the neural network. The major difference between QNN and ANN is that the QNN uses quantum neuron multi-level neuron with sigmoid activation function. The QNN is composed of multi-level hidden units, the learning rate μ for weigh adjusting is set to 0.05 training by MATLAB test, and the learning rate for quantum interval adjusting μ is set to 0.001. Synaptic weights are adjusted by minimizing quadratic error function with respect to particular weights, in training itself; jump-positions of multi-level hidden units adjusted also. Results are done with 750 samples for training phase and 150 samples for testing phase, performance of QNN classifier give the accuracy of (98.245 %) for training phase and accuracy of (97.992 %) for testing phase.

We compare this result with ANN classifiers, first one is that QNN classifier gave better accuracy for testing phase compared with ANN classifier (98.245 % compared with 95.214 %) and second issue is that QNN converged with less number of iterations and this indicates that the time required for QNN convergence is less than that of ANN. The reason is that ANN is unable to correctly assign class member ship to data samples belonging to regions of the feature space where there is overlapping among the classes. The reason for this is that FFNNs use sharp decision boundaries (due to crisp membership function) to partition the feature space. As a result, the outputs of trained ANNs cannot generally be interpreter as membership values. Also it can be found that QNN is more reliable than the ANN because QNN generates a more structured representation of the input data at the hidden layer than that of the ANN as QNN use multilevel hidden units, this is not surprising, given the fact that the jump-positions of the multi-level hidden units of the QNN are updated by minimizing some measure based on the class-conditional variances at the outputs of the hidden units. Another advantage is that QNN systems are using quantum neuron instead of traditional neuron which is often able to learn faster and require less number of neurons in the hidden layer which could lead to a smaller number of weights and reduction of the number of neurons in the hidden layer which could lead to smaller number of weights, or it can be said generally that this means reducing the total number of parameters (input weights, output weights, jump position) to be learned by almost half, as an assumption to the total number for hidden units in hidden layer.

5 Conclusion

In this paper, we have proposed a malware classifier system based on QNN trained with back-propagation. The proposed system extract malware sample and normalize them using API system calls and tri-gram model then transform to quantum form for dataset. The architecture of QNN has inherently built in fuzzy, hidden units of these networks to develop quantized representations of sample. In simulation results we show that QNN is superior than ANN with the accuracy of 98,245 % vs 95,214 %. This accuracy is still too low in practice so that we will try to improve this result in the future.

References

1. Abninder, L., Chris, E., Frederick, W.K., Steven, W., Paul, T.: Is the Brain a Quantum Computer? No. 30, pp. 593–603. Cognitive Science Society, University of Waterloo (2006)
2. Rumelhart, D.E., Hinton, G.E., Williams, R.J.: Learning representations by back-propagating errors. Nature **323**, 533–536 (1986). London
3. Abd Al-Majeed, G.H., Alisa, Z.T., Naji, H.S.: Data classification using quantum neural network. J. Eng. **20**, 1–15 (2014)
4. Hinton, G.E., Osindero, S., Teh, Y.W.: A fast learning algorithm for deepbelief nets. Neural Comput. **18**, 1527–1554 (2006). MIT Press

5. Dahl, G.E., Stokes, J.W., Deng, L., Yu, D.: Large-scale malware classification using random projections and neural networks (2013)
6. Purushothaman, G., Nicolas, B.: Quantum neural networks (QNNs): inherently fuzzy feedforward neural networks. IEEE Trans. Neural Netw. **2**, 1085–1090 (1996)
7. Jarernsri, L., Mitrpanont, A.S.: The realization of quantum complex-valued back-propagation neural network in pattern recognition problem. In: The 9th International Conference on Neural Information Processing (ICONIP'OZ), vol. 1 (2003)
8. Kolter, J.Z., Maloof, M.A.: Learning to detect and classify malicious executables in the wild. J. Mach. Learn. Res. **7**, 2721–2744 (2006)
9. Grover, L.K.: Quantum mechanics helps in searching for a needle in a haystack. Am. Phys. Soc. **79**(2), 325–328 (1997)
10. Schultz, M.G., Eskin, E., Zadok, E., Stolfo, S.J.: Data mining methods of detection of new malicious executables. In: Proceedings of the 2001 IEEE Symposium on Security and Privacy, pp. 38–49 (2001)
11. Idika, N., Mathur, A.P.: A survey of malware detection techniques. Technical Report, Purdue Univ (2007)
12. Tsai, X.Y., Chen, Y.J., Huang, H.C., Chuang, S.J., Hwang, R.C.: Quantum NN vs. NN in signal recognition. In: Proceedings of the Third IEEE International Conference on Information Technology and Applications (2005)

Managing Secure Personal Mobile Health Information

Chan Wai Chen, Mohd Azam Osman(✉), Zarul Fitri Zaaba, and Abdullah Zawawi Talib

School of Computer Sciences, Universiti Sains Malaysia, 11800 Pulau Pinang, Malaysia
cwchen.ucom11@student.usm.my, {azam,zarulfitri,azht}@usm.my

Abstract. Medical errors may cause serious public health problems and threaten the safety of the patients. Part of the errors is due to mistakes in the medical record or incomplete medical record which may trigger tragic consequences. In this paper, we present an application that manages securely personal health information on a mobile platform and keeps all the medical records of a patient in digital format. Patients are able to access their medical record at their convenience and the confidentiality of information is guaranteed. Patients are also able to share their personal health record with their respective doctor in a secure way. The application consists of several modules: incognito, access control, privacy control, authentication, encryption, multifactor authentication and emergency control. An anonymous database is created by removing all the identifier of a patient before the health record is stored in the database. This provides an extra layer of protection to the patient's privacy. In particular, our proposed application introduces the multifactor authentication and emergency control modules which provides a multi-layered defense authentication and emergency case handler respectively. Thus, the proposed application allows the patient to assess their records conveniently and securely, and helps them in emergency situations. As such, the application is suitable for cases involving large number of patients and emergency situations such as in Hajj healthcare management.

1 Introduction

Nowadays, health issue is one of the major concerns for everyone. According to the Global Status Report on Non Communicable Disease, 36 million people die each year due to non-communicable diseases such as cardiovascular disease, cancer, diabetes and chronic lung disease [1]. This constitutes almost two-thirds of the estimated 56 million deaths each year worldwide [1]. Research has proven that early diagnosis is the key to recovery. Therefore, having instantaneous and complete patient's medical history would help the clinicians to retrieve the patient's information and make better decisions. "Tens of thousands of people die every day, often for the simple reason that the parents, caregivers or health workers have insufficient information and knowledge that they need to save them" [2].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_38

Every year, more than two million Muslim pilgrims from all over the globe gather to perform the Hajj ritual in Mecca. This event creates an extraordinary density of people in many places around the holy city. Hot and humid weather during the day and night makes the situation more challenging for the pilgrims, especially to those over 60 years old. Risk factors that lead to the death of the pilgrims include high risk diseases, physical activities and immobility due to health condition, malnutrition and dehydration [3]. The main issue in such a high density crowd is how to manage, monitor and retrieve the pilgrim's health information in the event of a sudden circumstance such as an accident. In this situation, a medical officer should be able to retrieve instantly the medical record of a patient. As a solution, we propose a secure mobile application for managing health information which can be used to keep personal health information and allow this information to be accessed by medical practitioners when consent is given to them. This application can certainly be used as one of the solutions for Hajj healthcare management in Mecca.

This solution aims to reduce the medical slip-up due to inadequate information given to a medical practitioner. This application is designed by taking into consideration the individual information privacy. Therefore, it is integrated with a high level security of data protection of the patient's personal and medical information.

The next section will present the background of the existing applications and this is followed by the system design of the application. Finally, screenshots of the user interface of the application and challenges and limitations in implementing the application will be discussed.

2 Background

In this section, we present some related works on health information systems. Some existing applications may be similar in concept with the proposed application. Yet, these applications have some limitations which include the inability to fulfill the requirements of the users and medical practitioners.

My Medical [4] is a comprehensive record keeping application developed for patient's personal medical information. My Medical is one of the alternative replacements for unreliable paper records or various electronic systems that hold bits and pieces of patient medical history. With My Medical, most of the information that are important to the patient will be kept together in one place. My Medical provides extra information where a traditional personal health record (PHR) does not provide such as emergency contacts, health insurance and doctors' contact information. In addition, since the patient already has medical contacts in the address book, My Medical provides a function to import these data automatically with just one click. Users do not need to retype them all over again. My Medical provides a range of common test result templates for patients to track their test results such as blood pressure, cholesterol level, metabolic panels, blood sugar level, and CT scan result. Should the provided templates found to be insufficient, My Medical allows a user to construct and customize

patient's test results according to their need in order to keep track their medical information. My Medical is basically an application for keeping all the medical information on the mobile devices such as the tablet and mobile phone. Currently, My Medical supports iOS and Android platforms and can be downloaded from App Store and Google Play.

Hello Doctor [5] is a free mobile app that allows you to connect with registered doctors by sending and receiving messages anywhere and anytime. It's a preventive care application that gives medical advice and guidance to the users to organize and control their health as well as understand their medical records (paper or digital). Hello Doctor allows the patient to share the medical records with his doctor and takes notes on the medical records by themselves. Any issues concerning the patients' health and treatment could be recorded directly by the patients. The issues including the doses of the medication drug, side effects, and treatment options can be stored directly in the medical record so that the patient's medical history can be more useful and understandable. The development of Hello Doctor was initiated after its founder's partner was diagnosed with a brain tumor. Mayaan Cohen attended doctor's appointments with her partner, and noticed that the information and keeping hard copies of medical information from each specialist were difficult and tedious. Thus, she realized that the only viable solution was to develop a simple and easy application as a means to keep record pertaining to health information after an appointment with each doctor on hand at all times. Hello Doctor works on mobile devices such as iPhone, Samsung, Blackberry and HTC. The full-featured Hello Doctor is available on iOS platform that is currently intended for iPad. It keeps appointments' record that can only be accessed by patients who have a password to the information. Furthermore, the app also provides statistical information on new medications. Moreover, Hello Doctor is able to predict the probability of various treatments.

iTriage [6] was created by two emergency room physicians to help people answer health questions such as "What might I have?" and "Where could I go for treatment?". iTriage has been used by over 1,000,000 people as a symptom checker, doctor finder, or medical reference tool. iTriage helps users to diagnose and identify the possible disease, find the most suitable doctors who can treat them and check the disease conditions as well as be notified on the closest appropriate health care providers according to the current location or any location the users might choose.

Microsoft HealthVault [7] is a web-based platform developed by Microsoft which allows individual to store and maintain health and fitness information in one place. Microsoft Health Vault is the personal health record system which was initiated in October 2007. It can be used for both individuals and healthcare professionals for keeping track all health details from simple to complex health issues. Examples of health details are medication, health history, blood pressure, condition and illness, allergy and fitness. One of the advantageous features provided by HealthVault is the multiple individual's access to individual's health information. For instance, a mother can access and manage records for each of her children or a son could have access to his father's record in order to

Table 1. Summary of features of the existing applications.

Feature	MyMedical [4]	HelloDoctor [5]	iTriage [6]	HealthVault [7]
Platform	Mac OS/iOS	iOS	Web/ Mobile	Windows,Web/Mobile
Target User	Patient	Patient	Patient	Patient
Medical Record	Yes	Yes	Yes	Yes
Managed by Doctor	-	-	-	Yes

aid his father deal with medical issues. Another significant feature provided by Microsoft HealthVault is permitting most important health information available to the medical attendant during emergency situations. Users can access and interact with their health information record through the HealthVault website or through a mobile application which is connected to the HealthVault platform. Authorization of access to the account can be done through Windows Live ID, Facebook or a limited set of OpenID providers. When a user firstly uses a HealthVault application, the user is required to set and authorize the application to access to a specific set of data and make it visible to other users. The users could also share some parts of the data or the whole of their health record with other users such as doctor, spouse and parent.

Several mobile applications are for Hajj Health Management such as Hajj Management system [8], SmartHajj [9] and Hajj Guider [10]. One of the features provided by the Hajj Management system is the adding of notes on the health issues of the pilgrims so that the authority can be alerted when the pilgrims request for emergency help. SmartHajj provides online feature which allows a user to connect with the authority directly. If the user falls sick, he is able to forward a request directly to the medic on standby. Hajj Guider provides features such as synchronizing it with the wearable band, monitoring friends and family members, and health checking for group members.

Table 1 presents a summary of features of the existing applications. The features include the platform used, target users, medical record information, user management (including doctors). In our proposed application, multifactor authentication and emergency case handler are introduced. Multifactor authentication refers to a security system that requires more than one method of authentication from independent categories to verify the user's identity during the authorization process.

3 System Design and Implementation

The proposed application is a one stop center designed for individuals and doctors to store and update an individual medical record into an anonymous database as well as to retrieve personal information. It allows an individual to share their private health record with the respective doctor in a highly secure way both in normal and emergency situations. It consists of two main components namely a centralized server, and mobile and Web applications. It is designed to provide the following features:

- An anonymous database stores the patient's health records.
- The doctor is allowed access to patient's health records in a secure way.
- Access is granted to the doctor with the permission of a patient's family member when an emergency case happens.

3.1 System Architecture

Figure 1 shows the system architecture of the proposed application which consists of the client application (Web and mobile applications) that allows the user to access the application. The web application will serve as the main platform for all users whereas the mobile application is a portable application that allows patients to access the application and grant the access to a third party (e.g., a doctor) from their mobile devices. The personal health record centralized server (Centralized Server) stores the patient's information and medical records as well as controls the access level of the health records. The data transfer between the client side's application and the server is highly secure using client/server authentication and data encryption methods. The server uses the patient's master ID to link records with their respective patient accordingly in the anonymous database. Even if the records are exposed to unauthorized parties, it is very unlikely for them to identify the respective patient from the database. Therefore, privacy is guaranteed. The encrypted master ID that is used to link with the patient's identity of the health records is encrypted by patient's secret key which is generated locally on his personal device. Only the doctor who is authorized is able to retrieve the master ID and access the patient record.

3.2 System Modules Breakdown

The main modules proposed in the application are the Centralized Server (Consisting of Incognito, Access Control, Privacy Control and Encryption modules) and the Mobile and Web Applications (Consisting of Access Control, Encryption, Multifactor Authentication and Emergency Control modules).

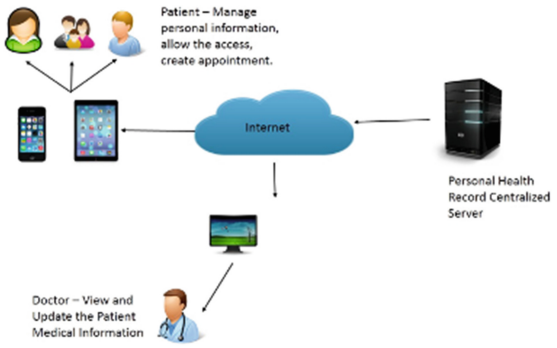


Fig. 1. System architecture diagram

Incognito module: This module is designed to protect the personal health record from intruders. The centralized server stores the records in incognito form. This module will discard all identifiers of the records (e.g., profile information) associated with the health record before the records are stored in the database. This technique is to ensure that a particular record cannot be associated with a specific identifiable individual. Thus, the intruder will not be able to determine which record belongs to any particular patient even though the intruder has managed to access the database. Each record will be associated with a unique Master ID created by the patient. The unique Master ID is generated after the patient's successful registration. The ID number is randomized and unique in such a way that no one is able to guess the identity of the patient from the patient's personal information. The patient does not need to remember this unique Master ID. The unique Master ID is appended to each of the health records during the new record creation process. Therefore, in the database, each of the records will only be associated with the unique Master ID. The Master ID is stored locally in the keychain of the mobile device. During the approval session, the patient will only reveal his/her master ID to the authorized doctor. The master ID is encrypted by the doctor's secret key and saved in the database.

Access Control module: This module is used to define the access control to the system based on the user's setting. Access Control module can be found on both the Centralized Server, and Mobile and Web Applications modules. Users need to register into the system before they can access to the application. The Access Control module in the Centralized Server is for the patients to control the access to the patient's personal health record, personal privacy control management data, and the personal sharing list. The detail of the privacy control and sharing list will be included in the privacy control module. After a successful fingerprint authentication process, patients are allowed to access to the application as well as retrieve their medical records. The Access Control module in the Mobile and Web Applications module controls the access to the application through the identification process of the user's identity by providing the user id, password and secret key. User can only access the records of the patient's information based on the setting in the patient's personal policy and list of shared records setting.

Privacy Control Module: This module is created to allow patients to control the access level of their health information in the application. Each patient is allowed to create and update the policies for accessing the health records by third parties (e.g., healthcare personnel or doctors who create the records). In the patient profile, the patient is able to set general access rules for healthcare professionals for accessing his health records. The patient is able to select one of the three types of access level namely low, medium and high. The doctor with low access level is not able to review the medical record, which mark as medium or high access level. The doctor with the highest access level granted by the patient will have full access to all of the records of the patient.

Authentication Module: This module is used to verify the identity of the patient and the doctor. During the registration process, a new patient profile is created in the system. A new unique Master ID and a secret key are generated locally on the mobile device. The patient's important information will be saved locally in the keychain of the mobile device. The keychain can be only accessed once the patient unlocks the phone with their fingerprint. As a result, the intruder will not be able to retrieve the private information from the patient's mobile device. The patients have two options in order to access the application, either by a common authentication process which is via the username and password or the fingerprint Touch ID authentication, a biometric authentication method which is more secure.

Encryption Module: Encryption Module is used to protect patient information from being stolen by an unauthorized party. The master ID that is the linker between the patient's profile and the medical records is encrypted by using the AES encryption. The master ID will only be revealed once the patient approved the doctor to access to his/her medical record. The master ID will be encrypted with the doctor's secret key. That means only the authorized doctor will be able to access to the patient's medical records.

Multifactor Authentication Module: A multifactor authentication module (MFA) is implemented on the Web Application. MFA is developed to provide a multi-layered defense authentication so that it will be more difficult for unauthorized party to gain access to the application. During the registration process, the doctor is required to create a unique username, password and secret key. Then, the doctor needs to take a photo. The photo will be presented to the face detection module for the recognition process. The application will be able to remember and recognize the identity of the doctor. During the login stage, the doctor needs to provide a valid user name and password. Both password and secret key is encoded using the bcrypt hashing algorithm [11] function before the login request is sent to the server for password verification. The reason for choosing this algorithm is due to its ability to calculate a huge amount of data in a short time. The hashed value is then compared with the values stored in the application database. If the values are valid then the doctor needs to undergo face recognition authentication to verify his/her identity before gaining access to the application. A third party cloud-based face recognition API is deployed to verify the face recognition in order to identify the doctor identity. Figure 2 shows the flow chart of the Multifactor authentication process.

Emergency Control Module: The emergency module is used when an emergency occurs. The patient can activate the emergency module by shaking the mobile device. Once the mobile device detects the motion, the patient is required to verify his/her identity before the activation of the Emergency Control module. If the verification process is successful, the current location will be identified through GPS and emergency alert will be activated. A notification will be send to the patient's family member and the doctor.

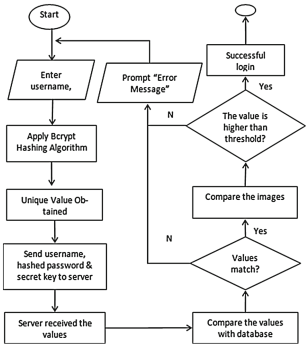


Fig. 2. Flow chart of multifactor authentication process

4 User Interfaces

Figures 3 and 4 show some sample screenshots of the user interface of the mobile application. Figure 3(a) shows the registration interface where the patient needs to provide a unique username and password. Figure 3(b) shows the login interface where the patient is able to login to the system by Touch ID authentication. The patient needs to upload a profile photo when using it for the first time. Figure 3(c) shows a successful login.

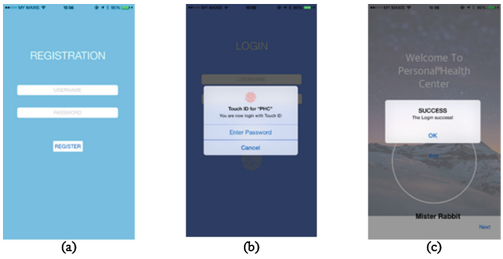


Fig. 3. Screenshots of the user interface of the mobile application interface (a) Registration page; (b) Login page; (c) A successful login

Figure 4(a) shows the emergency module activation, where the patient is able to activate the emergency module by shaking the device after launching the application. The patient needs to verify his/her identity before activation of the emergency module. The application will send the current location of the patient to his/her personal doctor and family members within 20 km radius. The emergency alert module (as shown in Fig. 4(b)) will trigger the application to send a notification to his/her personal doctor and family members.

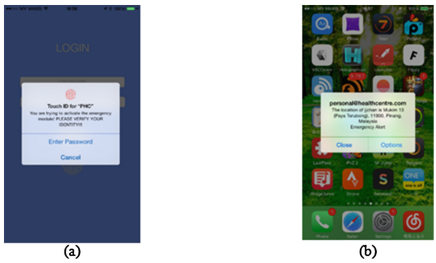


Fig. 4. User interface of the Emergency Control module (a) Activation; (b) Alert

5 Challenges and Limitations

During the design and development process of this project, many challenges have been faced, especially when designing the Graphical User Interface (GUI) for the iPhone. Designing the GUI including all the Table ViewController and CustomTableCell is not an easy task as both the Storyboard Interface Design and the implementation need to be combined by using programming code. Since the application includes many different fields and input, designing the database is also another challenge. Moreover, the mobile application needs to be compatible with the Web application. The data entered through the application must be able to retrieve information from the Web Application since both platforms share the same database.

The limitation of this application is that the Centralized Server does not have a secure socket layer. Therefore, data sent between the server and the client is vulnerable to eavesdropping. If an attacker is able to intercept all data being sent between a browser and a web server they can see and use that information. In order to prevent such attack, certificate authority (CA) must be purchased and implemented. Besides, the face recognition authentication is developed by using a cloud-based face recognition API which means that it may not be reliable since the authentication process is done by a third party. The authentication process will need to take some times to finish due to the transmission of image data to the server. The connection with the Centralized Server may not be stable because a free host provider hosts the server. The connection is slow and sometimes due to limit on the connection time, the connection may fail.

6 Conclusion

In the proposed application, the stored medical record is made anonymous so that the intruder will not be able to associate the records with any individual. The Master ID created by the patient is private and is only known by the patient and the authorized doctor. The advantages of the proposed application are that the database is anonymous, the patients have full control over their own health records and the confidentiality of any respective individual is secure.

The access security is further enhanced by the Multifactor Authentication module by integrating face detection and recognition process using the traditional method; username and password. The biometric authentication method is also implemented on the mobile application for an added protection to verify the identity of the user by using the fingerprint Touch ID. An emergency control module is also introduced to allow users to get help during an emergency situation.

In a nutshell, the proposed application is introduced to allow the patient to assess their records conveniently and in a secure manner. The application is suitable for cases involving a large number of patients and emergency situations such as in Hajj healthcare management.

Acknowledgement. We would like to acknowledge the contribution of the School of Computer Sciences, Universiti Sains Malaysia and the support of the Long-Term Research Grant Scheme (LRGS), 203.PTS.6728004 from the Malaysian Ministry of Higher Education.

References

1. Alwan, A.: Global Status Report on Non-communicable Diseases. World Health Organization, Switzerland (2010)
2. Thousands dying every day for lack of health information. World Confederation for Physical Therapy, January 2010. http://www.wcpt.org/node/33001
3. Gaffar, H.D., Achmadi, U.F., Syamsu, I.P.: Hajj health management focussing on the risk factor management. Int. J. Sci. Res. Publ. **3**, 1–9 (2013)
4. My Medical - The Personal Medical Record for You, The Patient. Hyrax Inc. http://mymedicalapp.com/
5. Hello Doctor. https://www.hellodoctor.co.za/
6. Find a Doctor, Symptom Checker, Conditions, Medications, Procedures and Hospitals. https://www.itriagehealth.com/
7. HealthVault Microsoft. https://www.healthvault.com/my/en
8. Binsalleeh, H., Mohammed, N., Sandhu, P.S., Aljumah, F., Fung, B.: Using RFID tags to improve pilgrimage management. In: IEEE International Conference on Innovations in Information Technology (IIT 2009), pp. 1–5 (2009)
9. SmartHajj. http://www.gistic.org/smart-hajj/
10. Hajj Guider, Pakistan Hajj Management. http://www.wareable.com/wearable-tech/hajj-guider-startup-will-help-pilgrims-get-to-mecca-1447
11. Provos, N., Maziéres, D., Sutton, T.J.: A future-adaptable password scheme. In: USENIX Annual Technical Conference, pp. 81–92 (1999)

Mobile Online Activity Recognition System Based on Smartphone Sensors

Dang-Nhac Lu[1(✉)], Thu-Trang Nguyen[1], Thi-Thu-Trang Ngo[2], Thi-Hau Nguyen[1], and Ha-Nam Nguyen[1]

[1] University of Engineering and Technology, Vietnam National University in Hanoi, Hanoi, Vietnam
{nhacld.di11,haunguyen,namnh}@vnu.edu.vn, trangnt84@gmail.com
[2] Posts and Telecommunications Institute of Technology, Ho Chi Minh City, Vietnam
trangntt1@ptit.edu.vn

Abstract. In this paper, we propose an efficient and flexible framework for activity recognition based on smartphone sensors, so called Mobile Online Activity Recognition System (MOARS). This system comprises data collection, training, activity recognition, and feedback monitoring. It allows users to put their smartphones in any position and at any direction. In our proposed framework, a set of power-based and frequency-based features is extracted from sensor data. Then, Random Forest, Naïve Bayes, K-Nearest Neighbor (KNN), and Support Vector Machine (SVM) classification algorithms are deployed for recognizing a set of user activities. Our framework dynamically takes into account real-time user feedbacks to increase the accuracy of activity prediction. This framework is able to apply for intelligent mobile applications. A number of experiments are carried out to show the high accuracy of MOARS in detecting user activities when walking or driving a motorbike.

Keywords: Activity Recognition · Mobile sensor · Online training model · Motorbike · User feedbacks

1 Introduction

Activity recognition plays an important role in context-aware applications such as security, healthcare, transportation, social networking, etc. In early time, most of researches analyze data that are collected from wearable sensors [1]. Nowadays, due to the emerging of smartphones, researchers switch their focus to the inference of user activities based on analyzing smartphone sensor data.

In the category of physical activity recognition, most of previous works focus on recognizing basic daily activities such as walking, standing, sitting, running, jumping, etc. [2]. Besides, a number of works tend to distinguish driving events, for example turning left, turning right, braking, etc. [3]. In this work, we concentrate on identifying the moving and stopping status of users while walking or driving a vehicle that allows estimating the total energy consumption of travelers.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_39

An activity recognition framework is normally composed of the following steps: First, the recorded sensor signal is broken into a number of windows of predefined size. Then, a set of features representing the general characteristics of each window is extracted. Finally, a classification algorithm is used to predict user activities. In fact, there are various approaches to deploy this framework. Most previous researches use classifiers that are trained offline [12]. As a result, the training process is static. These systems may be not adapted to new users. Example, some users might walk quite slower than others do. Therefore, the online training strategy is more preferable [8]. Our system follows this later strategy and also uses the assistant feedback from users to empirically increase the prediction accuracy.

There are various approaches for extracting features from sensor data, such as time domain features, frequency domain features, wavelet features, and power based features [14]. More recently, Khan et al. [11] and Han et al. [15] show that using a group of tilt invariant features, power based features, is not only computationally feasible in real-time systems but also robust to the rotation of accelerometer and reasonably accurate. In this work, we combine these features with a number of frequency-based features to represent the characteristics of each data window.

Indeed, very few existing driving event recognition methods have been tested for primitive vehicles such as motorbikes that are very popular in developing countries [14]. We thus carried out a number of experiments to verify the effectiveness of our activity recognition framework while travelers are not only walking but also using a motorbike.

The main contributions of this work: proposing a set of suitable features; suggesting a strategy that detecting user activities in a dynamic mode which utilizing real-time user feedback to increase the prediction accuracy; proposing an activity recognition framework that can be performed solely on smartphone, it thus allows preserving user privacy; carrying out experiments on primitive vehicles i.e. motorbike; using data collection strategy that allows users smartphones can be placed at any direction and any position during their trips [13].

2 Mobile Online Activity Recognition System (MOARS)

The details of our proposed mobile online activity recognition system (MOARS) are described in Fig. 1. This system is composed of three main modules: First, the Labeled data collector module is responsible for collecting smartphone sensor data for each predefined activity of each volunteer. Such data is then preprocessed, and extracted into a set of representative features that serves as inputs for the Training module. In the Training module, the selected classification algorithm is trained by using the computed features of each predefined activity. Based on the trained knowledge, the real time activity of users is detected by the Monitoring module. The interfaces of various modules of MOARS are presented in Fig. 2:

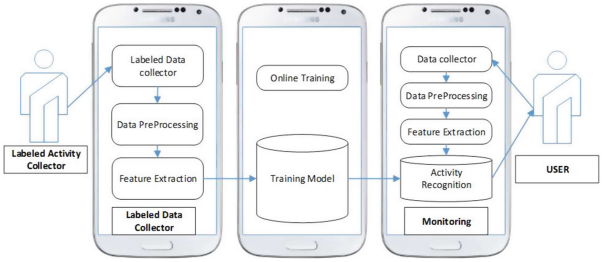


Fig. 1. Mobile online activity recognition system framework.

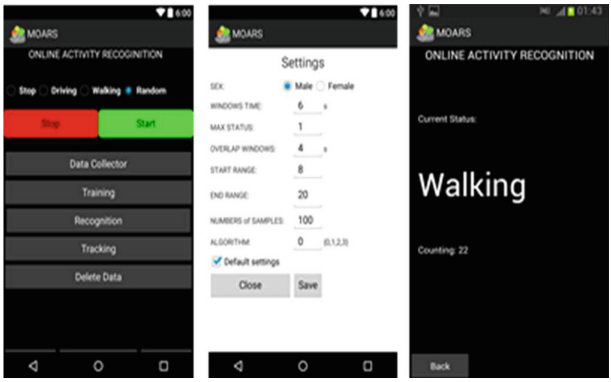


Fig. 2. The interfaces of the data collector and the monitoring module.

2.1 Data Preprocessing

In the present day, smartphones incorporated with many applications are pervasively used in the world. User information achieved from smartphone sensors, especially accelerometer, is very useful for activity recognition. However, users might put their smartphones on their pocket, handbag, or in their hands, etc. while moving. As a result, the orientation of smartphones will be frequently changed. Consequently, smartphone sensor data is noisy. An approach to solve this issue is to transform accelerometer data from the smartphone coordinate system (Fig. 3a) to the Earth coordinate system (Fig. 3b) by relying on the additional data collected from magnetometer, and gyroscope sensors. For the details of this transformation, we refer readers to the work of Premerlani and Bizard [2]. With smartphone sensor, there are three main method to solve streaming data as Explicit Segmentation, Time Base Windowing, sensor Event Based Windowing, it was shown in [13]. The Time Base Windowing is suitable [7] when integrated with hi-pass and low-pass filter into MOARS.

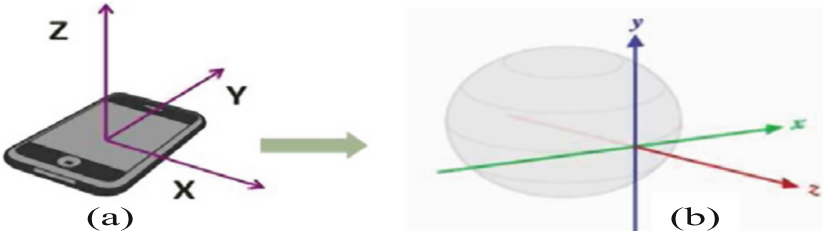


Fig. 3. (a) The smartphone coordinate (b) the Earth coordinates.

2.2 Feature Extraction

Indeed, the amount of raw data collected from smartphone sensors is huge. Hence, directly analyzing such data would require a lot of either time or memory space. A popular approach to deal with this issue is extracting certain important features from such data. Several types of features have been proposed, for instance time-domain features, frequency-domain features, wavelet, etc. [10]. Selecting suitable features would lead to the increment of the prediction accuracy. Therefore, we propose method combined extracted features from time-domain features and frequency-domain features as Fig. 4 below:

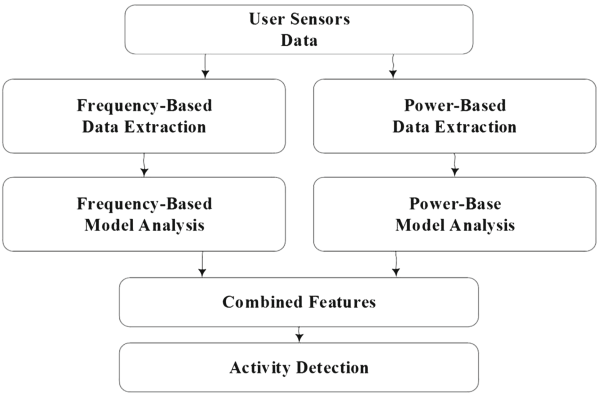


Fig. 4. The proposed set of features.

For real time activity recognition applications, Han et al. [15] suggested to use a set of tilt features that are that are accelerometer energy, Hjorth mobility, Hjorth complexity. In order to improve the quality of feature set, we combine these power based features with several frequency-based features that are the correlation between the horizontal and vertical acceleration time series, the number of signal peaks, the total energy of the peaks, and the velocity variation in each window:

$$e_v = \int_{t=t_0}^{T+t_o} |a_v \mid \, dt, \qquad e_h = \int_{t=t_0}^{T+t_o} |a_h \mid \, dt \tag{1}$$

where a_v, a_h are respectively vertical and horizontal acceleration, and T is the interval of integration in which $a_v(t) = a_z(t)$, $a_h(t) = \sqrt{a_x^2(t) + a_y^2(t)}$ Hjorth Mobility and Hjorth Complexity were introduced by Hjorth in 1970 to analyze time series data [8]. For a time series a_1, a_2,...,a_N, the Hjorth Mobility is computed as $\sqrt{\frac{M_1}{TP}}$ and the Hjorth Complexity is defined as $\sqrt{\frac{TP.M_2}{M_1.M_1}}$ where $TP = \frac{1}{N}\sum_{i=1}^{n} a_i$, $M_1 = \frac{1}{N}\sum_{i=1}^{n} d_i$, $M_2 = \frac{1}{N}\sum_{i=1}^{n} (d_i - d_{i-1})^2$, and $d_i = a_i$ - a_{i-1}. Applying these formulas on the vertical and horizontal acceleration time series, we obtain the resulting vertical and horizontal mobility and complexity. The correlation between the horizontal and vertical acceleration time series is calculated by the following formulas:

$$R = \frac{\sum_{i=1}^{N} \left(a_h^i - \bar{a}_h\right)\left(a_v^i - \bar{a}_v\right)}{\sqrt{\sum_{i=1}^{N} \left(a_h^i - \bar{a}_h\right)^2 \sum_{i=1}^{N} \left(a_v^i - \bar{a}_v\right)^2}} where \quad \bar{a}_x = \frac{1}{N}\sum_{i=1}^{N} a_x^i \tag{2}$$

The number of signal peaks in a window of transformed accelerometer data is computed as the following:

$$a_{rms} = \sqrt{a_x^2 + a_y^2 + (a_z - g)^2}, \bar{a}_{rms} = \frac{1}{N}\sum_{i=1}^{N} a_{rms}^i, a_{thr} = k.\bar{a}_{rms} \tag{3}$$

where N is number of signal in the data window. With k = 1, $a_p = a_{rms} - a_{thr}$ Using the zero-crossing algorithm to count the number of peaks by:

$$d(n) = \left| \frac{\text{sgn}\left[a_p(n)\right] - \text{sgn}\left[a_p(n-1)\right]}{2} \right| \tag{4}$$

with n = 2,...,N ; sgn is the sign function.

$$P = round\left(\frac{1}{2}\sum_{n=1}^{N} d(n)\right) \tag{5}$$

The total energy of the peaks, the area of the peaks, is computed by:

$$S = \frac{1}{2}\sum_{n=2}^{N} \left[a_p(n-1) + a_p(n)\right] * \left[t(n) - t(n-1)\right] \tag{6}$$

where $a_p(n) > 0$.

The velocity variation in a window data is computed by the below formulas:

$$\Delta V_i|_{t1}^{t2} = \int_{t1}^{t2} a_i dt \tag{7}$$

with i = x, y ,z axis and window time [t_1,t_2]. In the frequency based, we compute Short Time Fourier Transform on window number n included N samples [x_n,x_{n+1},...,x_{n+N-1}] as:

$$X(n,k) = \sum_{m=0}^{N-1} x[n+m].w[m].\exp\left(-j\left(2\pi/N\right).k.m\right) \tag{8}$$

with k = 0, 1, ..., N-1 and w[m] as window function. The Energy of M coefficient Fourier is computed by below formula:

$$E_M = \sum_{m=1}^{M} |X(m)|^2 = \sum_{m=1}^{M} X(m).X^*(m) \tag{9}$$

The average Energy of Z axis also computed as:

$$\bar{E} = \frac{2\sum_{m=2}^{N/2} |X(m)|^2}{N} \tag{10}$$

Finally, the Entropy is computed by below formulas:

$$H = -\sum_{m=1}^{N} p_m \log_2\left(p_m\right) \tag{11}$$

with $p_m = \frac{|X(m)|}{\sum_{m=1}^{N} |X(m)|}$

With that equations, we was computed ten features and combined to the suitable set of features for MOARS (Fig. 4)

2.3 The Online Training Model

Classification is an important step in the activity recognition process. The most commonly used classifiers are decision tree, support vector machine (SVM), K-nearest neighbor (KNN) and Naive Bayes [6]. In fact, a classifier first needs to be trained by using labeled activity database (training data). There are two training approaches: the offline and the online mode. In the online mode, the training is performed on smartphones. On the other hand, the offline training is done in advance, usually on a local machine. Most of existing studies use the offline training method [12]. The reason is to reduce computational cost on smartphones. Nonetheless, the modern smartphones have much better computational capacity. This improvement of smartphones allows us to implement the online training in our MOARS.

2.4 Real Time Feedback

In a traditional activity recognition framework, the training data set is usually fixed and prepared in advance. Since each user might have different habits, for instance one might walk faster than others do. As a consequence, the prediction accuracy might fall down when the system is used for new users. Gomes et al. [8] proposed an idea to incrementally update the training data set by utilizing real-time feedbacks from users. We implement this approach in MOARS as following: For each data window, the system provides the corresponding activity prediction. Then, user confirms the correctness of the result. If the prediction is correct, the data window labeled by the predicted activity is then added to the training dataset. The challenge here is to efficiently process such additional information.

3 Experiments and Results

3.1 Testing Environment

We implemented MOARS our mobile online activity recognition system on the Android from 4.0 to 5.0 platform. The labeled activity database is collected by 20 subjects when walking or driving a motorbike. They freely carry a Samsung galaxy S4, Quad-core 1.6 GHz Cortex-A15 processor, 2 GB of Ram, 2600 mAh battery, Android 4.2.2 Jelly Bean. The set of activities for recognition is stopping, walking, driving.

3.2 Data Collection

In our scenario, data is collected from 3 types of sensors: acceleration sensor, gyroscope sensor and magnetic sensor. Each sensor returns three values corresponding to x, y, and z coordinates. The raw data stream is first cut out 2 s at the beginning, and 4 s at the end as these time periods are usually redundant. Then, it is split into a number of windows of 6 s; the overlapping time is one second. We collected 3000 samples for each activity from 20 subjects.

3.3 The Accuracy of Activity Detection

We used four classification algorithms employed in the WEKA tool to predict the travelers status: Random Forest, KNN, Naive Bayes, SVM. In each case, the default setting was used. For evaluating the accuracy of each classification algorithm, we used 10-fold cross validation. In the first scenario, our experiment processing with 20 subject and collected 3000 sample data per activity, then we did offline classifying on smartphone and get result as:

We used four classification algorithms employed in the WEKA tool to predict the status of travelers: Random Forest, KNN, Naive Bayes, SVM. In each case, the default setting was used. For evaluating the accuracy of each classification

Table 1. The prediction accuracy of MOARS offline without using user feedback

	Random forest	KNN	Naive Bayes	SVM
Stopping	89.15 %	80.00 %	70.00 %	66.45 %
Driving	80.65 %	78.17 %	62.00 %	52.00 %
Walking	90.79 %	88.00 %	81.00 %	67.50 %
Average	86.86 %	82.06 %	71.00 %	61.98 %

algorithm, we used 10-fold cross validation. In the first scenario, MOARS predicts user activities offline classifying without using user feedbacks. The obtained results are depicted in Table 1.

In the second scenario, MOARS predicts user activities Online and utilize user feedbacks to enhance its prediction accuracy. The prediction results are illustrated in Table 2.

Table 2. The prediction accuracy of MOARS when using user feedbacks

	Random forest	KNN	Naïve Bayes	SVM
Stopping	94.64 %	85.20 %	80.00 %	68.00 %
Driving	89.00 %	80.15 %	72.00 %	69.50 %
Walking	90.30 %	86.00 %	85.00 %	73.15 %
Average	91.31 %	83.78 %	79.00 %	70.22 %

Tables 1 and 2 show that the prediction accuracy of MOARS is critically improved when user feedbacks are used in most of cases. Especially for the case of predicting Driving activity by SVM, the accuracy is increased by 17.5 %. These improvements highlight the effectiveness of the online activity recognition strategy used by MOARS. Moreover, the results in Tables 1 and 2 also indicate that Random Forest is more suitable for our MOARS framework since it always leads to higher accuracy as comparing with the other classifiers, i.e. KNN, Naive Bayes, and SVM. Table 2 shows the accuracy for detecting the current activity of smartphone users by MOARS can be up to 94.64 % when Random Forest classifier is used. In fact, the accuracy for detecting Driving activity is lower than that of others. The reason is due to misinterpreting some similar patterns such as slowly driving and stopping. Note that, our MOARS framework allows detecting the current user activity in the condition that their smartphones can be put at any position and in any direction. In previous studies, the framework proposed by Berchtold et al. [9] could achieve 97 % accuracy when smartphones were fixed, but only 63 % accuracy when smartphones were randomly placed, the real time knowledge-driven [4] activity recognition framework proposed by Chetty et al. obtained only 80 % accuracy [6].

3.4 Proccessing Time

Taking into account the real time feedback from users allows increasing the accuracy. However, it usually requires additional time for processing such information. Table 3 shows the average time to detect each type of activity by MOARS. As we can see, Random Forest also spend least time for detecting user activity as comparing to KNN, Naïve Bayes, and SVM.

Table 3. The average detecting time required by MOARS (seconds)

	Random forest	KNN	Naïve Bayes	SVM
Stopping	2	2	3	4
Driving	1	3	2	2
Walking	3	3	3	4
Average	2	2.7	2.7	3.3

4 Conclusion

In this paper, we proposed a flexible framework, called MOARS, for detecting current user activity when smartphones are randomly placed in any position and at any direction. In addition, our proposed framework also utilizes real-time feedbacks from users to increase the prediction accuracy. In the experiments, MOARS can achieve on average 91.31 % accuracy for detecting three predefined activities, i.e. stopping, walking, and driving a motorbike. Moreover, Random Forest classifier is a promising one for our framework. In the future, we are planning to further improve the current framework to either increase the activity prediction accuracy or reduce the processing time.

References

1. Bao, L., Intille, S.S.: Activity recognition from user-annotated acceleration data. In: Ferscha, A., Mattern, F. (eds.) Pervasive 2004. LNCS, vol. 3001, pp. 1–17. Springer, Heidelberg (2004). doi:10.1007/978-3-540-24646-6_1
2. Premerlani, W., Bizard, P.: Direction cosine matrix imu: theory, Technical report, 17 May 2009
3. Saiprasert, C., Pholprasit, T., Thajchayapong, S.: Detection of Driving Events using Sensory Data on Smartphone. Int. J. Intell. Transp. Syst. Res. 1–12 (2015)
4. Okeyo, G., Chen, L., Wang, H., Sterritt, R.: Dynamic sensor data segmentation for real-time knowledge-driven activity recognition. Pervasive Mob. Comput. **10**, 155–172 (2014)
5. Chetty, G., White, M., Akther, F.: Smart phone based data mining for human activity recognition. Procedia Comput. Sci. **46**, 1181–1187 (2015)
6. Laguna, J.O., Olaya, A.G., Borrajo, D.: A dynamic sliding window approach for activity recognition. In: Konstan, J.A., Conejo, R., Marzo, J.L., Oliver, N. (eds.) UMAP 2011. LNCS, vol. 6787, pp. 219–230. Springer, Heidelberg (2011). doi:10.1007/978-3-642-22362-4_19

7. Gomes, J.B., Krishnaswamy, S., Gaber, M.M., Sousa, P.A.C., Menasalvas, E.: MARS: a personalised mobile activity recognition system. In: 2012 IEEE 13th International Conference on Mobile Data Management (MDM), ISBN: 978-1-4673-1796-2, pp. 316–319 (2012)
8. Berchtold, M., Budde, M., Gordon, D., Schmidtke, H., Beigl, M.: Actiserv: activity recognition service for mobile phones. In: 2010 International Symposium on Wearable Computers (ISWC), pp. 1–8, October 2010
9. Menshawy, M.E.L., Benharref, A., Serhani, M.: An automatic mobile-health based approach for EEG epileptic seizures detection. Expert Syst. Appl. **42**(20), 7157–7174 (2015)
10. Khan, M., Ahamed, S.I., Rahman, M., Smith, R.O.: A feature extraction method for real time human activity recognition on cell phones. In: RESNA-ICTA Conference, Toronto, ON, Canada, 5–8 June 2011
11. ur Rehman, H.M., Liew, C.S., Wah, T.Y., Shuja, J., Daghighi, B.: Mining personal data using smartphones and wearable devices: survey. Sensors 2015 **15**, 4430–4469 (2015). doi:10.3390/s150204430
12. Krishnan, N.C., Cook, D.J.: Activity recognition on streaming sensor data. Pervasive Mob. Comput. **10**, 138–154 (2014)
13. Preece, S.J., Goulermas, J.Y., Kenney, L.P.J., Howard, D.: A comparison of feature extraction methods for the classification of dynamic activities from accelerometer data. IEEE Trans. Biomed. Eng. **56**(3), 871–879 (2009)
14. Nguyen, T.-T., Nguyen, T.-H., Nguyen, H.-N., Nguyen, D.-N., Choi, G.S.: Detecting user status from smartphone sensor data. IJCC **4**, 28–30 (2016)
15. Han, C.W., Kang, S.J., Kim, N.S.: Implementation of HMM-based human activity recognition using single triaxial accelerometer. IEICE Trans. Fundam. **93**(7), 1379–1383 (2010)

Multi-criteria Path Planning for Disaster Relief: An Example Using the Flood Risk Map of Shalu District, Taiwan

Sheng-Wei Qiu[1], Yi-Chung Chen[1(✉)], Tsu-Chiang Lei[2], Hsin-Ping Wang[2], and Hsi-Min Chen[1]

[1] Department of Information Engineering and Computer Science, Feng Chia University, No. 100, Wenhua Road, Xitun District, Taichung City 407, Taiwan, R.O.C.
shadow3x3x3@gmail.com, {chenyic,hsiminc}@fcu.edu.tw

[2] Department of Urban Planning and Spatial Information, Feng Chia University, No. 100, Wenhua Road, Xitun District, Taichung City 407, Taiwan, R.O.C.
tclei@fcu.edu.tw, blue00857@hotmail.com

Abstract. Route planning is one of the most common location-based services today. Conventional route planning algorithms could only consider a single criterion, which cannot meet actual needs. Skyline path queries were therefore proposed for multi-criteria path queries. However, we discovered that most existing skyline path query algorithms are extremely slow, and while such slow speeds may be acceptable to the general public, they are unacceptable for disaster relief. We therefore proposed two means of improvement. Experiment simulations demonstrated the validity of our approach.

1 Introduction

Route planning is a crucial part of disaster relief systems. In the past, routes were determined as "good" based on single conditions, such as requiring the least time or being the shortest distance. An example is Google Map. However, in reality, people consider more than one condition when choosing a route. For instance, giving consideration to both distance and time would most likely produce better routes. Route planning disaster relief would encompass even more conditions, such as the grade and reliability of the road or its susceptibility to flooding. To take into account multiple conditions in route planning, Tian *et al.* [5] and Kriegel *et al.* [3] extended the concept of skylines [1,2,4] and proposed skyline paths. We use Fig. 1 and Table 1 to explain the concept of skyline paths. In undirected graph G, suppose there is origin node s and destination node t. There are multiple routes p starting from s and leading to t. Each road link of the routes possesses multiple dimensions, and summing the dimensions of the links, or edges, in each route gives the dimensions of the entire route, as shown in Table 1. For candidate routes *p1* and *p2*, the former is superior to the latter in both time and distance, so we say that *p1* dominates *p2*. Continuing on,

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_40

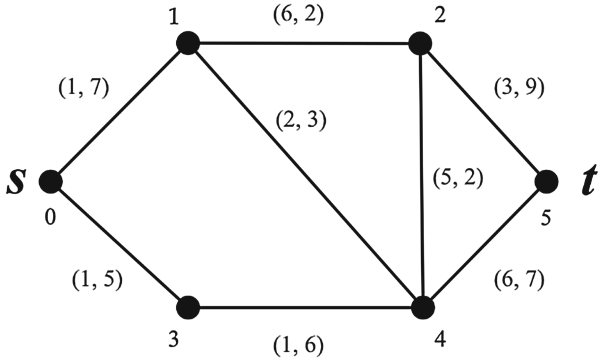


Fig. 1. Example for skyline paths

p3 requires less time but is longer in distance than *p1*, so the two cannot be compared, which means that both are skyline paths. After comparing all of the routes, the skyline paths are the routes that are not dominated by any other route (*p1* and *p3*). To identify skyline paths with multiple dimensions, we can only use the brute force approach, in which the sums of all of the dimensions of all the routes are calculated, or attempt to eliminate or limit certain dimensions, such as only taking into account distance, time, and risk of flooding while neglecting all other dimensions. Although these methods do get the job done, the dimensions are not compared at the same level, which means that the routes identified may not be very reliable. One well-known algorithm for identifying skyline paths is SkyPath, which stores the skyline path results or circumstances (the origin and destination) from each query it receives. The next time it receives a query with the same conditions, it will first compare the skyline paths stored in its database to reduce the number of nodes that needs to be visited. Although this approach is very straightforward and effective, it is not efficient in queries for disaster relief routes and may therefore delay disaster relief operations. We have found that some of the algorithm is in need of improvement. For instance, it does not produce a large number of dominating skyline paths, and the data structure design can incur considerable computational costs. We thus proposed a new algorithm in response to these shortcomings. The remainder of this paper is arranged as follows. Section 2 presents related works, Sect. 3 introduces the proposed algorithm and explains our simulation, Sect. 4 contains the conclusion.

2 Related Work

At present, the most well-known skyline path algorithms are those proposed by Tian *et al.* [5] and Kriegel *et al.* [3], which we introduce below.

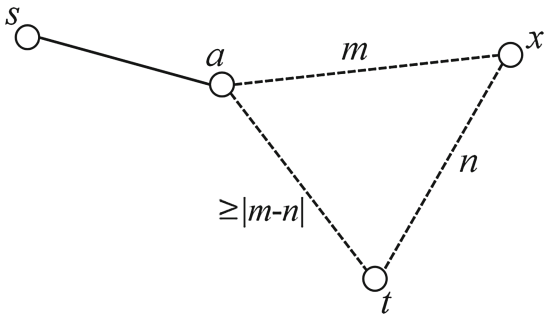


Fig. 2. Concept of multi-criteria path planning approach proposed by Kriegel *et al.* [3]

Table 1. Routes in Fig. 1 and their dimensions

id	Path	(Time, Distance)
p_1	0,1,4,5	(9, 17)
p_2	0,1,2,5	(10, 18)
p_3	0,3,4,5	(8, 18)
p_4	0,1,4,2,5	(11, 21)
p_5	0,3,4,2,5	(10, 22)
p_6	0,3,4,1,2,5	(13, 25)

2.1 Algorithm Developed by Tian *et al.* [5]

In the road network in the directed graph proposed by Tian *et al.* [5], each edge is designed with d dimensions. Queries are performed by summing the d dimensions of the edges in each route leading from a given origin to a given destination and then comparing the results. The approach proposed by the researcher is based on the Dijkstra algorithm. The path with the lowest sum for a certain dimension is first identified as the first result. The costs needed for each path not reaching the destination to reach the destination are then calculated. For example, given an origin s, destination t, and path $<s, a>$, which does not yet reach the destination, the cost required to go from a to t is first calculated, supposing that cost is X. If the sum of X and the d dimensions of $<s, a>$ is dominated by other paths, then none of the paths leading from s to t via a will be the solution to the multi-criteria problem.

2.2 Algorithm Developed by Kriegel *et al.* [3]

The method proposed by Kriegel *et al.* [3] differs from that presented by Tian *et al.* [5]. The former makes use of established reference nodes on the map rather than directly using the Dijkstra algorithm to derive the solution. In Fig. 2, for example, x is the reference node. With x being previously established, the minimum cost of the d dimensions of all the paths leading from x to t can be

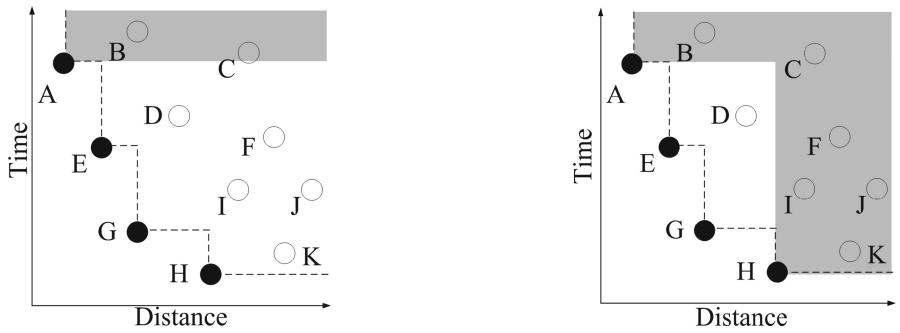


Fig. 3. The concept of dominance in advance using the minimum data points for each dimension

calculated beforehand. Suppose the minimum cost of getting from x to a is m and that of getting from x to t is n. Thus, from the triangle formed by a, x, and t on the map, it can derived that the cost of getting from a to t must be greater than or equal to $m - n$ (the difference between two sides of a triangle must be less than the other side). Furthermore, the difference between the costs of $<a, x>$ and $<x, t>$ is the minimum cost of $<a, t>$. Compared to the approach proposed by Tian *et al.* [3], this method derives the costs of getting from the reference node to the destination beforehand. As a result, when queries are made, the algorithm only needs to calculate the cost of the shorter $<a, x>$ and does not need to calculate the cost of $<a, t>$, which reduces the computation time considerably.

3 Algorithm

The purpose of this study was to apply multi-criteria path planning to obtain routes for disaster relief operations. We therefore focused on skyline paths and endeavored to make improvements on the design of the existing SkyPath algorithm to ensure query speed. SkyPath first searchers for the shortest route, so this route must be a skyline path and can thus be used to quickly eliminate nodes that do not need to be visited when searching for new paths. However, the shortest route can only dominate a small portion of the nodes in terms of the dimension of distance (such as the rectangle dominated by node A in Fig. 3(a)). Finding the minimum skyline path for each dimension means that more nodes can be dominated beforehand (such as the rectangle dominated by nodes A and H in Fig. 3(b)) and reducing the number of unnecessary comparisons. This approach is particularly useful for the planning of routes for disaster relief, which involves relatively more factors. While this forms a high-dimension skyline path problem, the proposed approach enables us to identify more skyline paths at the beginning.

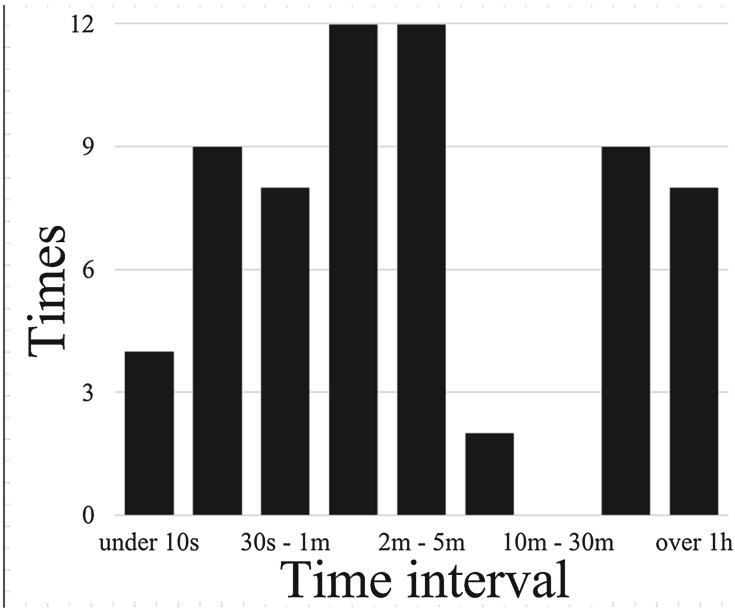


Fig. 4. Simulation results

In addition to making improvements regarding dominance, this study also attended to the data structure in SkyPath. Originally, SkyPath used arrays to process the skyline paths stored from previous queries. However, as the number of nodes visited and the number of paths found increase, so does the number of arrays. To avoid this circumstance, we propose the use of hashing to store routes in place of arrays. The key to the hash function contains the origin and destination of the route, and the key value is the dimension value of the route. This reduces the cost of visiting arrays as well as the time costs of computation.

4 Simulations

With the two improvements above, we performed a simple simulation experiment to verify their feasibility. The field of the experiment was the road network of Shalu District, Taiwan, and its flood risk map. Note that this location was chosen because it has seen frequent flooding in recent years and because it has a comprehensive network of flood sensors. Given the opportunity, our algorithm can be directly applied to it.

We conducted a total of 60 experiments. In each experiment, we simulated the behavior of the public and selected new origins and destinations. The results are presented in Fig. 4. As can be seen, 47 of the experiments were completed with 5 minutes, occupying 78% of all the experiments. This demonstrates the validity of the proposed method. As for the queries requiring more than an hour

to process, their origins and destinations were quite far from each other, which may have resulted in too many possible skyline paths and thus a greater number of dominance checks, thereby taking considerable amounts of time.

5 Conclusion

Skyline path queries have become a significant application in daily life; however, existing query methods are time consuming and thus cannot be applied to path planning for disaster relief. This study therefore proposed two improvements to accelerate queries. Experiment simulations demonstrated the validity of the proposed approach. We will continue our efforts in developing novel algorithms to accelerate skyline path queries, including (1) incorporating index structures to serve as the algorithm foundation and (2) adding more criteria to enhance algorithm efficiency.

Acknowledgement. This work was supported in part by the Ministry of Science and Technology of Taiwan, R.O.C., under Contracts MOST 105-2119-M-035-002, and MOST 105-2634-E-035-001.

References

1. Borzsonyi, S., Kossmann, D., Stocker, K.: The skyline operator. In: Proceedings on ICDE, pp. 235–254 (2001)
2. Chen, Y.C., Lee, C.: The σ-neighborhood skyline queries. Inf. Sci. **322**(11), 92–114 (2015)
3. Kriegel, H.P., Renz, M., Schubert, M.: Route skyline queries: a multi-preference path planning approach. In: Proceeding on ICDE, pp. 261–272 (2010)
4. Papadias, D., Tao, Y., Fu, G., Seeger, B.: An optimal and progressive algorithm for skyline queries. In: Proceeding on SIGMOD (2003)
5. Tian, Y., Lee, K.C.K., Lee, W.C.: Finding skyline paths in road networks. In: Proceeding on SIGSPATIAL GIS, pp. 444–447 (2009)

Multi-feature Based Similarity Among Entries on Media Portals

Thi Hoi Nguyen[1], Dinh Que Tran[2], Gia Manh Dam[1], and Manh Hung Nguyen[2,3(✉)]

[1] Vietnam Commercial University, Hanoi, Vietnam
hoint2002@gmail.com, damgiamanh@gmail.com
[2] Posts and Telecommunications Institute of Technology (PTIT), Hanoi, Vietnam
tdque@yahoo.com, nmh.nguyenmanhhung@gmail.com
[3] UMI UMMISCO 209 (IRD/UPMC), Hanoi, Vietnam

Abstract. Similar measures play an important role in information processing and have been widely investigated in computer science. With the exploration of social media such as Youtube, Wikipedia, Facebook etc., a huge number of entries have been posted on these portals. They are often described by means of short text or sets of words. Discovering similar entries based on such texts has become challenges in constructing information searching or filtering engines and attracted several research interests. In this paper, we firstly introduce a model of entries posted on media or entertainment portals, which is based on their features composed of title, category, tags, and content. Then, we present a novel similar measure among entries that incorporates their features. The experimental results show the superiority of our incorporation similarity measure compared with the other ones.

Keywords: Similar measure · Text · Entry · Social media · Feature

1 Introduction

With the exploration of entertainment, news or media portals and the social networks, a huge number of entries have been posted on these portals. Such entries could be a video clip, an image, a text, or a combination of them. They are often included short text or sets of words to describe viewpoints, comments and so on. Discovering similar entries based on such texts has become challenges in constructing information searching or filtering engines and attracted several research interests.

D. Lin [5] proposed a model to measure the similarity between any two objects based on information-theoretic approach. Nguyen and Nguyen [6] introduced a general model for measuring the similarity between objects based on their attributes. In these models, the similarity on each attribute is defined with different natures and types of values. On text similarity, the current researches could be classified into two main groups. The first approach considers the text

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_41

similarity based on their semantic. It means that texts are compared with each others based on their semantic similarity. And the semantic is then based on WordNet in the work given by Buscaldi et al. [2], or Lee et al. [4], or based on ontology such as Nguyen and Tran [7,11].

The second approach makes use of static techniques for modeling the text similarity. Texts are then compared based on the lexical structure of words, sentences; and/or based on the statistic of words in texts. Researches on similarity based on static include the work given Buscaldi et al. [1], Finkel et al. [3], Proisl et al. [8], Sultan et al. [10]. However, most of these models consider only the text body to estimate their similarity. They lack of investigating additional information such as the tags, category, title and keywords, which may contribute greatly to modeling the similarity of entities.

In this paper, we firstly introduce a model of entries posted on media or entertainment portals. Our model not only considers the text content of entries, but also takes into account their features of title (caption), category, and tags of entry to estimate the similarity. Secondly, based on the model, we propose an incorporation similarity measure that aggregates feature-based similarities into an overall similarity among entries. Experiments will be performed to validate and evaluate the performance of our model compared with other ones.

The paper is organized as follows. Section 2 presents the model of entries and their similarity measure based on similarities of features. Section 3 describes experiments to evaluate the proposed model. Section 4 is the conclusion and perspectives.

2 A Model of Entries and Similarity Measure

The general model is presented in Fig. 1. It takes the two entries as input data and the output is the estimated similarity between the two entered entries. Inside the model, there are three main steps:

- **Step 1:** Modeling entries
- **Step 2:** Estimating the similarity on each entry features after preprocessing text such as eliminating stop words etc.
- **Step 3:** Aggregating the similarity on the whole entry from their similarities on features.

These steps will be described in detail in the next sections.

2.1 Modeling Entries

Without loss of generality, we assume that:

- An entertainment portal contains a set of entries $\{1, 2...m\}$.
- An entry is described by their features such as title (caption), category, tags, and content. Each entry i has n features, denote $(f_1^i, f_2^i, ..f_n^i)$. For the purpose of this paper, we focus on describing four features of entries in our experiment:

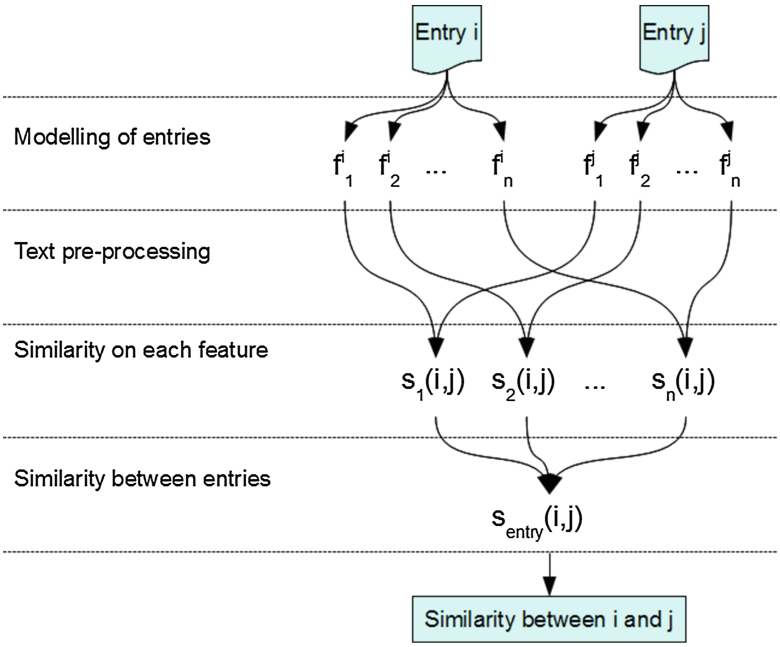


Fig. 1. Overview of the model

- *Title* of entry i, noted as f^i_{tit}: It could be an entitle of a brief sentence. Note that in the case of image, a title is considered as its caption if it has not any explicit title on the portal.
- *Content* of entry i, noted as f^i_{con}: An entry could be a video clip, an image, a text, or a combination of them. However, in this model, we consider only the textual parts in the value of entry features. Other media such as video, images etc., are out of scope of this paper. In the case that the content of an entry has no text, we considered as there is no available data for this feature.
- *Tags* of entry i, noted as f^i_{tag}: On the entertainment portal, each entry could be tagged to a set of tags. Each tag is an independent word or expression.
- *Category* of entry i, noted as f^i_{cat}: On the entertainment portal, each entry is always assigned to at least one category. Each category is represented by an independent word or expression.

In such a model, an entry is considered as a set of features and only their textual values are considered. And then the problem of estimating the similarity among entries becomes the computation of the similarity among texts or among sets of expressions.

2.2 Similarity on Each Feature

As only textual value of feature is considered, we need to take some steps of pre-processing the textual value before comparing of them. We distinguish two kinds of textual value of feature:

- First, the feature value is already in form of a set of expressions, such as the value of feature *tag*, and *category*. Their similarity is resulted to considered among sets of expressions.
- Second, the feature value is in form of a general text, such as the value of feature *content*. Their similarity is considered among texts.

In the case of feature being *title*, it is generally is a sentence. In spite of its short length, this feature is also considered as a text.

Feature value is a set of expressions. Since the feature value is in the form of a set of textual expressions, their similarity is defined as follows:

Suppose that $A_1 = (a_1^1, a_1^2, ...a_1^m)$, $A_2 = (a_2^1, a_2^2, ...a_2^n)$ are two sets of expressions or strings, where m, n are the sizes of the set A_1 and A_2, respectively. Let v be the size of the intersection of A_1 and A_2. The similarity between A_1 and A_2 is defined by the formula:

$$s_{exp}(A_1, A_2) = \frac{2 * \mid A_1 \cap A_2 \mid}{\mid A_1 \mid + \mid A_2 \mid} = \frac{2 * v}{m + n} \tag{1}$$

It is clear that all possible values of $s_{exp}(A_1, A_2)$ are lied in the interval $[0, 1]$. This formula could be applied to all considered features of an entry because these features are considered as a set of expressions. Suppose that $i = (f_1^i, f_2^i, ..f_n^i)$, $j = (f_1^j, f_2^j, ..f_n^j)$ are two entries represented by their features. Let consider the feature k whose value is a set of expressions. The similarity between entries i and j on the feature k is defined by the formula:

$$s_k(i, j) = s_{exp}(f_k^i, f_k^j) \tag{2}$$

where f_k^i, f_k^j are the expression values of the feature k of the two entries i and j. For example, the similarity on feature *tags* and *category* of two entries are given as follows:

$$s_{tag}(i, j) = s_{exp}(f_{tag}^i, f_{tag}^j) \tag{3}$$

$$s_{cat}(i, j) = s_{exp}(f_{cat}^i, f_{cat}^j) \tag{4}$$

Feature value is a text. In this case, the problem becomes the estimation the similarity between two texts. We could apply the technique TF-IDF (Term Frequency - Inverse Document Frequency) [9] to characterize the texts, which are used in many statistic-based models such as Buscaldi et al. [1], Finkel et al. [3]. In our work, TF-IDF is also used to estimate the similarity between two features of text value as follows:

– Extract the feature value (a text) into a set of n-gram $t^1 = (g_1^1, g_2^1, ...g_n^1)$ and $t^2 = (g_1^2, g_2^2, ...g_m^2)$
– Calculate the TF-IDF of each n-gram in the text. Then represent the feature value by a vector whose each element is a pair $<$ *n-gram, td-idf* $>$: $v^1 = (< g_1^1, v_1^1 >, < g_2^1, v_2^1 >, ... < g_n^1, v_n^1 >)$ and $v^2 = (< g_1^2, v_1^2 >, < g_2^2, v_1^2 >, ... < g_m^2, v_m^2 >)$
– Calculate the distance between the two vectors:

$$D(v^1, v^2) = \frac{1}{N} \sum_{1}^{N} d_k \tag{5}$$

where N is the number of different n-grams considered in both $t^1 \cup t^2$, d_k is the distance on each element $< g_i^1, v_i^1 >$ of v^1 (or element $< g_j^2, v_j^2 >$ of v^2, respectively):

- If there is an element $< g_l^2, v_l^2 >$ of v^2 (or element $< g_l^1, v_l^1 >$ of v^l, respectively) such that $g_l^2 = g_i^1$, then:

$$d_k = \frac{\mid v_i^1 - v_l^2 \mid}{max(v_i^1, v_l^2)} \tag{6}$$

- Otherwise, $d_k = 1$.

– It is clear that the value of $D(v^1, v^2)$ is in the interval $[0, 1]$. Similarity between the two features is then:

$$s_{txt}(t^1, t^2) = 1 - D(v^1, v^2) \tag{7}$$

For example, similarity between *title* and *content* of two entries i and j is as follows:

$$s_{tit}(i, j) = s_{txt}(f_{tit}^i, f_{tit}^j) \tag{8}$$

$$s_{con}(i, j) = s_{txt}(f_{con}^i, f_{con}^j) \tag{9}$$

2.3 Similarity Between Two Entries

Once the similarities between two entries on each feature are estimated, the similarity between two entries is then computed by a weighted average aggregation of the similarity on all considered features as follows.

Suppose that $i = (f_1^i, f_2^i, ..f_n^i)$, $j = (f_1^j, f_2^j, ..f_n^j)$ are two entries represented by their features. The similarity between entries i and j on all considered features is defined by the formula:

$$s_{entry}(i, j) = \sum_{k=1}^{n} w_k * s_k(i, j) \tag{10}$$

where $s_k(i, j)$ is the similarity on feature k of entries i and j; w_k is the weight of the feature k such that:

$$\sum_{k=1}^{n} w_k = 1 \tag{11}$$

The more this similarity is closed to 1, the more the two entries are similar. And vice versa, the more this similarity is closed to 0, the less the two entries are similar.

3 Experimental Evaluation

This section first describes the construction of sample set and then presents the experiments and evaluation results.

3.1 Construction of Sample Set

Each sample is constructed as follows:

- Each sample contains three entries collected from a source (Youtube, CNN News). These entries are called as entry A, B, and C.
- We ask a number of selected people to answer the question: Which entry, B or C, is more similar to entry A than the other one?
- Then, we compare the number of people who chooses B, and that of people who chooses C. If the number of answers B is greater than that of C, then the *value* of this sample is 1. It means that the entry B is more similar to the entry A than C is. In the contrary, if the number of the answer C is greater than that of B, then the *value* of this sample is 2. It means that the entry C is more similar than the entry A than B is. If the number of the answer B and C are not significantly different, this sample will be removed from the sample set.

For instance, let consider an sample with three entries extracted (on July 18th 2015) from Youtube as described in Table 1: Each entry is presented in a row of the table, each feature is presented in a column.

We then use entry A to compare with the two remain entries, by asking a group of 9 voluntary members to answer a question in our survey:

- In comparing to the entry A, among two entries B and C, which entry is more similar to entry 1 than the other? Designed answer is a choice of B or C.

Table 1. Data of three entries extracted from Youtube

Entry	Title	Category	Tags	Content
A	Top 30 Goals World Cup 2014	Sports	Worldcup, Football	no text
B	Top 10 Goals: 2014 FIFA World Cup Brazi [Official]	Sports	Worldcup, Football, Brazil, FIFA	no text
C	The Speech that Made Obama President	Education	Obama, President speech	no text

Table 2. Data collected from human behaviour on entries from Youtube

Question	Majority choice	Minority choice
1	9 (for entry B)	0 (for entry C)

Table 3. Sample constructed from three sources

Source	Number of samples
CNN News	100
Fox News	100
YouTube	100
Total	**300**

The number of people answered each question is then counted. The the principle of vote by majority is applied. The results are presented in Table 2.

These empirical results indicate that, entry B and A are more similar than the entries C and A. Therefore, the *value* of this sample is set to 1.

After this step, we have a set of samples. We could use several sources of sample and save them in several sets of sample. In experiments, we use three sources of sample with the size of each sample set is described in the Table. 3.

3.2 Method of Experiment and Evaluation

The experiment is performed as follows:

- For each sample, we use model proposed in this paper to estimate the similarity between the entry B and A, and that between entry C and A.
- If B is more similar to A than C is, then the *result* of this sample is 1. In the contrary, If C is more similar to A than B is, then the *result* of this sample is 2.
- We then compare the *result* and the *value* of each sample. If they are identical, we increase the variable *number of correct sample* by 1.

In order to evaluate the results, we make use of the correct ratio (CR) of the model over the given sample set which is calculated as follows:

$$CR = \frac{\text{number of correct sample}}{\text{total of sample}} * 100\,\% \tag{12}$$

The more the CR value is closed to 100 %, the more the model is correct. We expect that the obtained value of CR is high as much as possible.

3.3 Experiment and Evaluation

Choose weights for features. The objective of this experiment is to find the best weights of feature (w_1, w_2, w_3, w_4). So the scenario to run the experiment is as follows:

Table 4. Correct ratio CR (%) and corresponding weight of each feature

Source	Title only	Content only	Tags only	Category only
CNN News	69	74	77	31
Fox News	32	82	62	31
YouTube	72	-	62	26
Average CR (%)	**57.67**	**78.00**	**67.00**	**29.33**
Normalized	**0.25**	**0.34**	**0.29**	**0.12**

– Test over all samples each time for only feature *title*, *content*, *tags*, *category* of the entry by setting the respective weight of the test feature to 1, all other feature weight to 0 for each time. Observe the *correct ratio* CR.
– The more the single feature brings the higher *correct ratio* CR, the more the feature is important. So the weight of the feature is also higher.

The results are presented in the Table. 4. The important level ratio among features is $(w_1 : w_2 : w_3 : w_4) = (0.25 : 0.34 : 0.29 : 0.12)$. So we choose the corresponding weight $(0.25, 0.35, 0.30, 0.10)$ for all next experiments.

Comparing with other Models. This experiment will evaluate the proposed model by comparing it with some recent models. Two models of Buscaldi et al. [1] and Finkel et al. [3] are utilized since both of them are based on n-grams to split text into expressions, and use statistic method to measure the similarity among texts. Table 5 summaries some main technical features of these models compared with ours.

Table 5. Technical features of the three models

Model	Model of Finkel et al. [3]	Model of Buscaldi et al. [1]	Our model
Considering of entry feature	only text body (content)	only text body (content)	Title, category, tags, content
Method	statistic-based	statistic-based	statistic-based
Split text	n-grams (1-gram)	n-grams (1-gram)	n-grams (1-gram)

The scenario of experiment is as follows:

– Use the same set of 200 samples constructed from two sources (CNN News, and Fox News) as the input for all three models. The samples from Youtube are not used because they have no text in the body (content) of entry. Meanwhile the model of Buscaldi et al. [1] and Finkel et al. [3] are mainly based on the body text of entries to estimate their similarity.

Table 6. Correct ratio CR (%) in comparing with other models

Model	Model of Finkel et al. [3]	Model of Buscaldi et al. [1]	Our model
CNN News	74	71	96
Fox News	82	89	96
Average CR (%)	**78.00**	**80.00**	**96.00**

- Run the model of Finkel et al. [3] and Buscaldi et al. [1]: Use only the content (body text) of entry to estimate the similarity among entries. And then, observe the *correct ratio* CR.
- Test over all samples on the whole entry with our model: Use the whole four features of entry with weight $(w_1 = 0.25, w_2 = 0.35, w_3 = 0.30, w_4 = 0.10)$ to estimate the similarity among entries. And then, observe the *correct ratio* CR.
- Then compare the obtained results.

The results are presented in the Table 6. They indicate that our model is significantly better the model of Buscaldi et al. [1] and Finkel et al. [3], regarding the given samples. As the models of Buscaldi et al. [1] and Finkel et al. [3] consider only the text body to estimate their similarity, meanwhile our model takes also the title, tags, and category of the texts into account. These results confirm our argumentation that it could be better if several features of texts are taken into account instead of only the text body to estimate their similarity.

4 Conclusions

This paper presented a model of entries posted on media or entertainment portals, and estimating the similarity among these entries. The computation is based on the four features of entry, including title, category, tags, and content. The experimental results show the superior of our model compared with other ones. However, some issues should be considered furthermore. Currently, we are considering how to take the semantic of text into account to compare expressions. These research results will be presented in our future work.

References

1. Buscaldi, D., Rosso, P., Gomez-Soriano, J.M., Sanchis, E.: Answering questions with an n-gram based passage retrieval engine. J. Intell. Inf. Syst. **34**(2), 113–134 (2010)
2. Buscaldi, D., Le Roux, J., Garca Flores, J.J., Popescu, A.: Lipn-core: semantic text similarity using n-grams, wordnet, syntactic analysis, esa and information retrieval based features (2013)

3. Finkel, J.R., Grenager, T., Manning, C.: Incorporating non-local information into information extraction systems by gibbs sampling. In: Proceedings of the 43rd Annual Meeting on Association for Computational Linguistics, ACL 2005, Stroudsburg, PA, USA, pp. 363–370. Association for Computational Linguistics (2005)
4. Lee, M.C., Chang, J.W., Hsieh, T.C.: A grammar-based semantic similarity algorithm for natural language sentences. Sci. World J. **2014**, 17 (2014)
5. Lin, D.: An information-theoretic definition of similarity. In: Proceedings of the 15th International Conference on Machine Learning, pp. 296–304. Morgan Kaufmann, San Francisco (1998)
6. Nguyen, M.H., Nguyen, T.H.: A general model for similarity measurement between objects. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **6**(2), 235–239 (2015)
7. Nguyen, M.H., Tran, D.Q.: A semantic similarity measure between sentences. SE Asian J. Sci. **3**(1), 63–75 (2014)
8. Proisl, T., Evert, S., Greiner, P., Kabashi, B.: Robust semantic similarity at multiple levels using maximum weight matching. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, pp. 532–540. Association for Computational Linguistics and Dublin City University, August 2014
9. Salton, G., McGill, M.J.: Introduction to Modern Information Retrieval. McGraw-Hill, Inc., New York (1986)
10. Sultan, M.A., Bethard, S., Sumner, T.: Sentence similarity from word alignment. In: Proceedings of the 8th International Workshop on Semantic Evaluation (SemEval 2014), Dublin, Ireland, pp. 241–246. Association for Computational Linguistics and Dublin City University, August 2014
11. Tran, D.Q., Nguyen, M.H.: A mathematical model for semantic similarity measures. SE Asian J. Sci. **1**(1), 32–45 (2012)

MyEpiPal: Mobile Application for Managing, Monitoring and Predicting Epilepsy Patient

Nur Ayuni Marzuki(✉), Wahidah Husain, and Amirah Mohamed Shahiri

School of Computer Sciences, Universiti Sains Malaysia, USM,
11800 USM, Penang, Malaysia
{nayuni.ucom12,ams14_com044}@student.usm.my, wahidah@usm.my
http://www.usm.my

Abstract. An epileptic seizure can be defined as a disturbance of consciousness, motor function, sensation, emotion, and behavior. The seizures can affect almost all aspects of life, such as lifestyle, health, education, and employment. Without a proper tool, epilepsy patient and caregiver finds difficulty in managing and monitoring epilepsy. The caregivers' lifestyle is also affected when the patient strongly dependent on them. By using current technology, which is a mobile phone, MyEpiPal application will be implemented to solve these problems. MyEpiPal will act as a self-management and monitoring tool that can support the epilepsy patient and the caregiver. This application also will enable them to monitor side-effects and effectiveness of antiepileptic medicine, predict the possibility of seizure attack, and improve the quality of life. A fully functional mobile application will be a supportive tool for monitoring and managing epilepsy patient.

1 Introduction

Epilepsy is a neurological disorder that produces sudden surges of electrical energy in the brain. These episodes called seizures, may alter a person's awareness, movements, or actions for a few seconds or minutes. A patient who is diagnosed with epilepsy will suffer epileptic seizures that repeatedly occur regardless of place and time. Other than that, epilepsy can cause nerve cell activity in the brain becomes damaged, causing convulsion, unusual behavior, loss of consciousness, and also can lead to sudden, unexpected death of someone with epilepsy (SUDEP) [1].

According to Consensus Guidelines on the Management of Epilepsy 2010, there are two types of epilepsy treatment in Malaysia which is surgery and medicine [2]. But, not all patients can undergo surgery and survive like a normal person. Besides, not all of them are getting better after using one or two different medicines. They still suffer from epileptic seizure which needs daily observation. From the doctors' perspectives, they have problems with compliance towards medicine. Often, the doctors are faced with the dilemma of deciding the best treatment because the patient does not record the health conditions and medicine correctly. There is a problem where the caregiver needs to monitor the patient's

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_42

condition 24 h a day. This situation limits their movement to do other daily activities.

Thus, MyEpiPal, a mobile-based application is proposed which will help epilepsy patient and/or the caregiver to monitor and manage epilepsy. The development of MyEpiPal can be useful to help the epilepsy patient to improve their quality of life. The application is also highly beneficial to the caregiver and the doctor as well as the patient. The objectives of the proposed system are:

- To develop a mobile application as a self-management tool to support epilepsy patient and caregiver in managing epilepsy patient.
- To reduce the level of the dependence of epilepsy patient on the caregiver.
- To improve management of epilepsy by helping clinicians in making a diagnosis of epilepsy based on the monitoring data.

2 Background Study and Related Work

According to World Health Organization (WHO) records, epilepsy affects about 50 million people worldwide [3]. Nearly 80 % are found in developing countries, including Malaysia and approximately 198,903 Malaysians were diagnosed with epilepsy in the year 2004 [4]. Epilepsy is a neurological disease that can be divided into idiopathic, symptomatic and cryptogenic. Idiopathic epilepsy has no cause, no associated clinical signs, normal brain imaging, and normal electroencephalogram (EEG) background. While symptomatic epilepsy has known causes and cryptogenic have an underlying cause that cannot be documented objectively. Thus, cryptogenic epilepsy is more likely to be symptomatic than idiopathic. This type of epilepsy can cause epileptic seizures that can be classified into absence, cyclonic, colonic, tonic, tonic-colonic, and atonics [4].

In Malaysia, epilepsy is still a taboo topic. People still consider this disease as a condition due to mythic causes, such as possession [5]. People who believe this theory is often would see shamans or traditional healers to cure this disease. The shaman will use traditional herbal medicine to cure epilepsy. It is very different compared to modern treatment that uses anti-epileptic drugs (AEDs) or surgery to cure epilepsy. But both treatments either traditional or modern are not able to cure epilepsy completely and only can control seizures. In modern ways, history taking is the main diagnostic tool. Caregiver or the patient will be asked to record their seizures in epilepsy diary given by the hospital to support the diagnosis of the patient. Epilepsy patient needs to record the frequency of seizures attack, and medication intake.

Since the increasing of technologies nowadays, numerous innovative research and development have been done in order to improve the management of epilepsy. Medical technologies such as EEG and neuro imaging have already been used in hospitals in order to diagnose and identify epilepsy. In the context of monitoring and managing epilepsy patient for their daily routine, it is still a fresh and new topic to be researched in Malaysia. But, there are several research conducted in other countries regarding this topic to fulfill the epilepsy patient's desire based

on the current technologies which is smart phones. So, by proposing MyEpiPal as a self-management mobile application, this app can help the epilepsy patient and the caregiver to fulfill their requirements.

There are many health related conditions may have indicators that reflect a measurable factor such as motion or lack of motion. An accelerometer type sensor is appropriate to monitor health conditions such as sleep apnea, Sudden Infant Death Syndrome (SIDS), health conditions, and other health condition related to motion [6]. For example, epilepsy patient who may experience with falling down may be sensed by using an **accelerometer** that positioned on the body. So, accelerometers can be used to sense the elderly body movement and range in level intensity, such as sitting, standing, walking, shaking, and etc. Other than that, by using accelerometer, it senses the position in the absence of motion.

Other than that, a **microphone** which is a popular sensor also can be used to sense sound and may be appropriate for health conditions involving the emission of sound such as snoring. This sensor also can be used to sense the volume, frequency, pattern or any other aspect of the sound relevant to the snoring condition. The sensor will monitor the breathing pattern of patient. A microphone also can be used to sense the volume, frequency, pattern or any other aspect of the sound and can be used in heart disease where it sense patterns of heartbeat [6].

Lastly, **Global Positioning System (GPS)** sensor is really helping a patient who suffers from a neurological disease such as Dementia and Alzheimer [7]. This kind of disease should be constantly monitored. For example, an application that embedded with an accelerometer and GPS sensor, which check for possible falls of the patient and immediately notifying a predetermined person for such an incident [8]. There is also another application that able to trace Dementia and Alzheimer patients when they are outdoors by using GPS sensors. The caregiver can discover their exact position since GPS can help in allowing the device to localize itself [9].

3 Proposed System

MyEpiPal is an android-based mobile application which is developed to monitor and manage epilepsy patient. The aim of this project is to improve the quality of life for the patient and caregiver. By using MyEpiPal, it helps people with epilepsy by managing their seizure and medication taken records, provide reminders for taking medicine and medical appointment, predict a seizure, detect a seizure using motion and voice, alert the caregiver when a seizure occurs, trigger sound to alert people surroundings, and etc. MyEpiPal also provide information about the patient, epilepsy and first-aid step by step when a seizure occurs that could be useful for people who less concern about epilepsy.

There are four components in MyEpiPal which are seizure management, medicine and appointment management, seizure detection, and seizure prediction as describe in Table 1. The first function allows the patient or caregiver to manage their seizure event. Next, the patient or the caregiver can manage medicine and appointment by adding, updating and viewing medication intake

Table 1. Descriptions of MyEpiPal components

Components	Description
Seizure management	– Manage patient and emergency contacts information
	– Record a seizure video as an evidence that could beneficial during medical consultation
	– Manage details of seizure including additional information such as side-effects
	– Provide data visualization for data seizure using graft
	– Export seizure data recorded to clinicians
Medicine and Appointment Management	– Manage details of medicine, including side-effects of medication
	– Manage medical appointment details
	– Set reminders for medication intake and medical appointment
	– Monitor the effectiveness of antiepileptic medicine
	– Export medication data recorded to clinicians
Seizure Detection	– Detect motion when a seizure happens with accelerometer sensor
	– Detect the voice of the patient when having a seizure using microphone sensor
	– Sending emergency SMS with GPS location to alert caregivers when a seizure happens
	– Triggers sound to alert people surroundings
	– Produce seizure detection data report for clinical and forensic purposes
	– Contact nearest hospital for some cases that require an ambulance
Seizure Prediction	– Predict the possibility of a seizure attack using patient's current condition (mood, other disease, alcohol intake, sleep quality, and blood glucose level)

and medical appointment. Then, the seizure and medicine data recorded can be exported to the clinician. The patient or caregiver also can schedule reminders for medication intake and next medical appointment. Besides, the patient or

caregiver can manage seizure detection either using motion or voice to detect the seizure. An emergency SMS will be sent to the caregiver and a sound to alert surroundings will be triggered when a seizure happens.

The patient or the caregiver also can predict a seizure based on the patient's current conditions.

3.1 System Architecture

The Fig. 1 below is the system architecture of MyEpiPal. MyEpiPal is developed by using a client-server concept. Both patient and caregiver's smart phones must be connected to the Internet and GPS in order to ensure MyEpiPal application can work properly and can access the server. There are three users involved in this proposed architecture which are patient, caregiver and clinician. The patient and caregiver represent as authorized users that have the ability to use MyEpiPal application to monitor and manage epilepsy. MyEpiPal will manage all data recorded in SQLite Database and MySQL Database. Data inserted into the database can be accessed in the both patient and caregiver's smart phone.

3.2 System Flowchart

Figure 2a is a system flow chart for Seizure management is used to help epilepsy patient or the caregiver to manage seizures. This component would be beneficial for a clinician to monitor the frequency of the seizures and further diagnosis. From the figure above, the patient or the caregiver can manage a seizure event by adding and updating a seizure event details. There are several information regarding the seizure events that must be included such as type of seizure, time

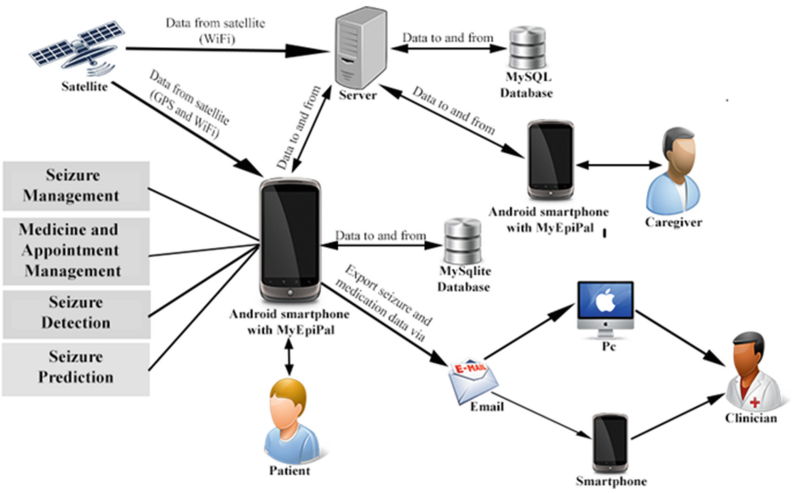
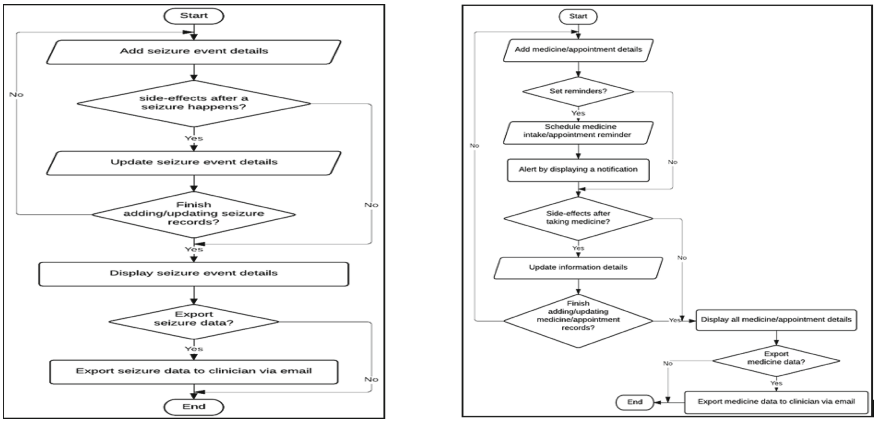


Fig. 1. System architecture

of a seizure, and for how long the seizure goes. If the patient got any side-effects after having a seizure, the patient or the caregiver must update the seizure event data. The data recorded can be viewed in data visualization and the patient or caregiver can export the data recorded in pdf format to clinician via email.

Figure 2b is a flowchart for Medicine and Appointment Management Component. This component is used to help epilepsy patient or the caregiver to manage medication intake and medical appointment. At the same time, this component allows the patient or caregiver to schedule reminders for medication intake and next medical appointment. This component would be beneficial for the patient, caregiver and clinician monitor the effectiveness of medication intake towards epileptic seizures. This component helps the clinician in decision making whether to continue the medicine, change medication's dosage or suggest new medicine. The medicine data also can be viewed and exported in pdf format via email to the clinician.

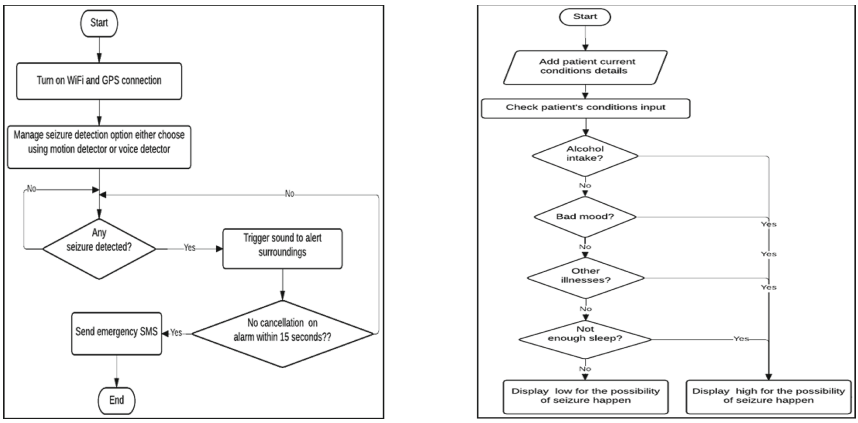


(a) Seizure Management (b) Medicine & Appointment Management

Fig. 2. (a) Seizure Management (b) Medicine & Appointment Management

Figure 3a is Seizure detection component describes the process for detecting any seizures when happen. There are two types of seizure detection used which are motion detector and voice detector. The motion detector is used to detect unusual movement of the patient that can be considered as a seizure. The voice detector is used to detect the voice of the patient in case of seizure attack. This voice detector highly recommended to be used while the patient sleeps. When a seizure is detected, an alarm sound will be generated to alert surroundings and within the given time, if there is no cancellation have been made by the patient, an emergency SMS will be sent to the caregiver.

The last component which is seizure prediction is used to predict any possibility of seizure attack based on the patient's current conditions. Figure 3b describes the flow of seizure prediction component. Based on the flowchart in



(a) Seizure Detection (b) Seizure Prediction

Fig. 3. (a) Seizure Detection (b) Seizure Prediction

the figure above, the patient or caregiver needs to input the patient's current conditions such as mood, quality of sleep, other illnesses, and alcohol use. Then, MyEpiPal will indicate the status of the possibility of seizure. MyEpiPal will inform the patient by displaying the status of the seizure to happen either high or low. By predicting the possibility of the seizure, the patient or caregiver can get ready for any unpredicted or bad situations. The algorithm used to predict the possibility of a seizure attack is a rule-based classifier as shown in Fig. 4.

```
BEGIN
CHECK mood for today
        IF mood is "Bored/Nervous/Sad"
        CHECK mood for yesterday
                IF mood is "Bored/Nervous/Sad"
                        THEN the possibility of seizure is high
CHECK alcohol intake for today
        IF alcohol intake is "200 grams or more"
                CHECK alcohol intake for yesterday
                        IF alcohol intake is "200 grams or more"
                        THEN the possibility of seizure is high
CHECK sleep quality for today
        IF sleep quality is "Lack of sleep"
                CHECK sleep quality for yesterday
                        IF sleep quality is "Lack of sleep"
                        THEN the possibility of seizure is high
CHECK other disease for today
        IF other disease is "Fever/Headache"
                CHECK other disease for yesterday
                        IF other disease is "Fever/Headache"
                        THEN the possibility of seizure is high
CHECK blood glucose for today
        IF blood glucose is "Low"
                CHECK blood glucose for yesterday
                        IF blood glucose is "Low"
                        THEN the possibility of seizure is high
CALCULATE the possibility of seizure based on these four parameters
END
```

Fig. 4. Prediction algorithm

4 System Interface

The screenshots of MyEpiPal app are shown below. Figure 5a is the main menu of the mobile app. Figure 5b and c are the sub-menus of seizure management component.



(a) MyEpiPal Main Menu

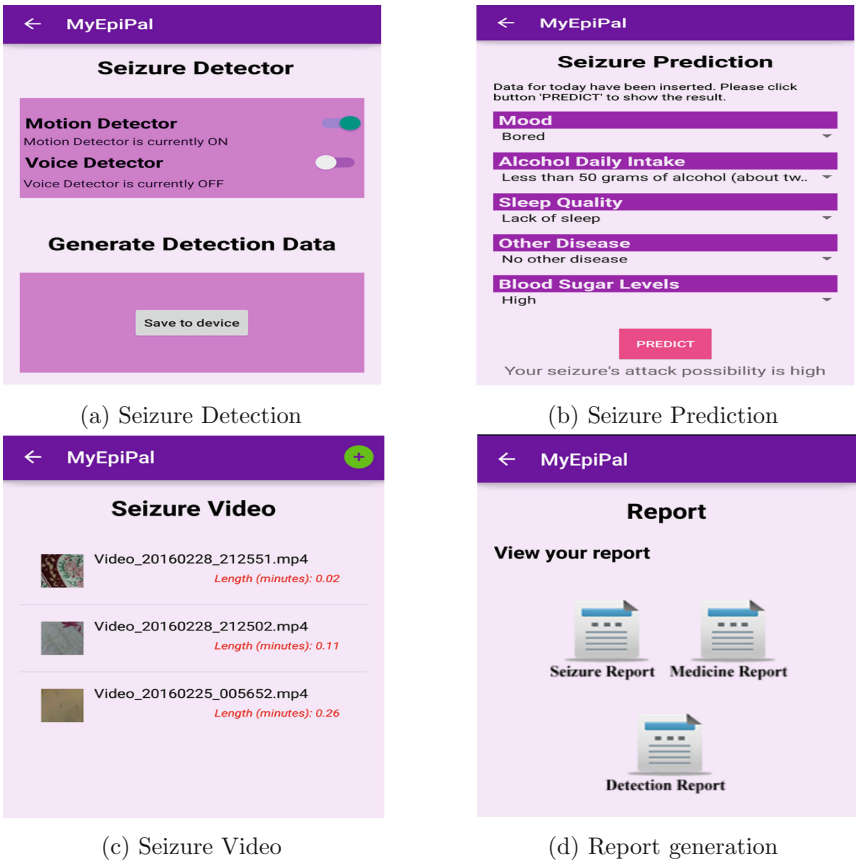
(b) Seizure Management

(c) Seizure attack

(d) Medication Notification

Fig. 5. (a) MyEpiPal Main Menu (b) Seizure Management (c) Seizure attack (d) Medication Notification

Figure 5d is medicine management sub-menu followed by seizure detector component as shown in Fig. 6a. Figure 6b is a seizure prediction sub-menu. Figure 6c is a list of recorded seizure attack videos. The last screen as shown in Fig. 6d is a different type of reports available that can be generated by the app, and the data can be displayed in pdf format for future references.



(a) Seizure Detection

(b) Seizure Prediction

(c) Seizure Video

(d) Report generation

Fig. 6. (a) Seizure Detection (b) Seizure Prediction (c) Seizure Video (d) Report generation

5 Conclusion

The analysis stage is a crucial part in system development to ensure all system requirements are included and show how these requirements will be accomplished. The related works based on the existing mobile applications for epilepsy have their own benefits and uniqueness. In order to get ideas to improve the proposed solution, all of these applications have been studied. Background study and the related works are really helping in order to make the proposed solution meet all requirements of the patient and the caregiver. With four components which are seizure management, medicine and appointment management, seizure detection, and seizure prediction, MyEpiPal will be able to support the need of the patients, caregivers and in order to monitor and manage epilepsy patient in an efficient manner and as well as improving the quality of their life.

Acknowledgments. We would like to show our gratitude to Universiti Sains Malaysia (USM) for supporting this research.

References

1. Annegers, J.F., Coan, S.P.: SUDEP: overview of definition and review of incidence data, International League against Epilepsy Millennium Meeting, Birmingham, United Kingdom (1999)
2. Consensus Guidelines on the Management of Epilepsy 2010, Epilepsy Council, Malaysian Society of Neurosciences (2010)
3. Epilepsy fact sheet, World Health Organization (2015). http://www.who.int/mediacentre/factsheets/fs999/en/. Accessed 29 September 2015
4. Sherrini, B.A.: What do you know about epilepsy? 15 March 2015. http://www.thestar.com.my/Lifestyle/Health/2015/03/15/What-do-you-know-about-epilepsy/. Accessed 1 October 2015
5. Teohari, V.-M., Ungureanu, C., Bui, T.V., Arends, J.B.A.M., Aarts, R.M.: Epilepsy seizure detection app for wearable technologies. In: European Conference on Ambient Intelligence, Eindhoven, The Netherlands (2014)
6. Rangenathan, L.N., Chinnadurai, S.A., Samivel, B., Kesavamurthy, B., Mehndiratta, M.M.: Application of mobile phones in epilepsy care. Int. J. Epilepsy **2**, 28–37 (2015)
7. Yang, J.: Toward physical activity diary: motion recognition using simple acceleration features with mobile phones. In: 1st International Workshop on Interactive Multimedia for Consumer Electronics, New York, USA (2009)
8. Ryan Burchfield, T., Venkatesan, S.: Accelerometer-based human abnormal movement detection in wireless sensor networks. In: 1st ACM SIGMOBILE International Workshop on Systems and Networking Support for Healthcare and Assisted Living Environment, New York, USA (2007)
9. Pigadas, V., Doukas, C., Plagianakos, V.P.: Enabling constant monitoring of chronic patient using android smart phones. In: 4th International Conference on Pervasive Technologies Related to Assistive Environments (2011). February 2016

New Block Ciphers for Wireless Moblile Netwoks

Pham Manh Tuan[1], Bac Do Thi[2(✉)], Minh Nguyen Hieu[3], and Nam Do Thanh[4]

[1] Posts and Telecommunications Institute of Technology, Hanoi, Vietnam
[2] University of Information and Communication Technology, Thai Nguyen, Vietnam
dtbac@ictu.edu.vn
[3] Academy of Cryptography Techniques, Hanoi, Vietnam
[4] Le Qui Don Technical University, Hanoi, Vietnam

Abstract. Wireless Mobile Networks (WMNs) play a fundamental role in the activities of each individual and organization. However, they have a number of limitations about hardware resources, processing capacity of the network devices. Hence, when integrating the cipher algorithms into the devices, the algorithms will have to be sure to reach high efficiency (fast processing speed and little resource expenses). This paper will demonstrate some results of constructing a new algorithm of cipher blocks based on the controlled element $F_{2/2}$. This is the controlled element which has been studied and proved to be suitable for developing the algorithms of block ciphers with highly integrated speed and efficiency on the hardware platform VLSI. Therefore, this algorithm will be sure to work effectively and suitable for transmitting secretly on WMNs. Furthermore, this paper will also propose the schedule which creates the on-the-fly key to counteract the attacks of the related key. Usage of this key schedule will still ensure high efficiency when integrating the cipher algorithms on WSNs.

1 Introduction

The popular block ciphers are usually developed with the aim of deploying effectively on both the software and the hardware, and as a result they are not suitable for the specific environment like WMNs. With respect to this network, there are some limitations about the processing capacity and the devices [1]. Consequently, the current trend of progressing block ciphers for WMNs is aiming to operate effectively about speed and be capable of integrating the hardware with high efficiency.

In fact, the encryption component of almost all applications in Wireless Sensor Networks requires quickly calculating time, which means the quantity and the complication of the encryption operations should maintain at a small level. The reason is, for the applications of this network, it merely needs to encrypt a small quantity of messages but requires the key to change frequently. One of the already-known trends, responding to the construction of the cipher algorithms

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_43

with high speed for WSNs, is using the algorithms based on DDP (Data Dependent Permutation) [2]. They are built up based on the controlled element (CE) $P_{2/1}$ and have been used as a main cipher element to design block ciphers such as CIKS-1 [2], CIKS-128 [2], Spectr-H64 [5], Cobra-S128 [2], Cobra-H64 [2] and Cobra-H128 [2]. However, the encryption algorithms based on DDP have potential weaknesses before the attacks based on linear cryptanalysis and differential cryptanalysis. This has been proved in the researches which are pointed out in the works. [2,5]

To eliminate the weaknesses of the encryption algorithms based on DDP, a number of encryption algorithms based on DDO (Data Dependent Operation) are established from elements $F_{2/1}$ or $F_{2/2}$, which were proposed on some researches. The encryption algorithms DDO-64 [2], DDO-128 [2], Cobra-F64a [2], Cobra-S128 [2], Eagle-64 [2], Eagle-128 [2] are the outstanding algorithms representing the class of algorithms which are built up from DDO. These algorithms have been proved to be appropriate to be implemented on the inexpensive and high-speed hardware. Yet, these algorithms only use simple key schedule so they are likely to be attacked by the related key [6,7,10].

For the encryption algorithms which are built up based on DDO (Data-driven Operations), to counteract the attacks of weak key (because they are designed according to the regulation of using simple key schedule), we can apply the method which creates the on-the-fly expansion key. The round keys are calculated parallel with the implementation of the previous encryption. In order to create the round key which applies a strong cipher procedure, it has to feature the high cascading effect when changing any bits of the secret key. Simultaneously, when applying the repeated procedure, its operation time needs to be smaller or equal to the operation time of the encryption round. The results must ensure not to slow down the algorithm when frequently changing the secret key (session key) and to only rise the resources negligibly when being implemented on FPGA.

Based on class CE $F_{2/2}$, this paper will demonstrate the results of constructing the algorithm of block ciphers with high efficiency TMN (including TMN64 and TMN128). Proofs, security estimation and efficiency of the developing algorithms will be summarized in the paper. Through the evaluation results, it is indicated that the algorithm proposed can entirely apply into practice to develop the communication applications on WSNs. The algorithms developed have the same structure, but use different strengths and weaknesses of the security and efficiency on the secret key and data's size. Choosing the size of the data block which is processed dissimilarly (64 or 128 bit) and selecting the length of different secret key allow them to apply into the networks with limited resources (these networks usually use the light algorithms of blocks ciphers) or the network environments which request the warranty of high security. The remaining of the paper will be presented with the structure: Sect. 2 will demonstrate the results of constructing the algorithm TMN; Sects. 3 and 4 will display security estimating and performance of the algorithm on FPGA; and the conclusion.

2 Developing the High-Speed Algorithm of Block Ciphers

This part will suggest the new algorithms of block ciphers, which are build up based on the element $F_{2/2}$. The algorithms are established with the same structure, in which the design of the right branch of processing data is completely new to the already-known architectures. To ensure the ability to counteract the attacks based on the weak key (the announced algorithms [6,7,10] based on the basis CSPN all apply the simple key schedule), the algorithms will be designed using the regulations of creating the on-the-fly round key. Thanks to these constructing regulations, the proposed algorithms will have high speed and performance on FPGA (proved in Sect. 4). The proposed algorithms are signalled as: $^{(2/2)}$**TMN64** and $^{(2/2)}$**TMN128**. These are the algorithms which are developed based on the element $F_{2/2}$. For convenience, in the process of describing them, we will only use signals **TMN64** and **TMN128**.

2.1 Describing the Algorithms

TMN64 and **TMN128** are the algorithms progressed with a view to operating effectively on the platform of FPGA. They are the algorithms of block ciphers whose block size is 64 or 128 bits and secret key size is 128 or 256 bits, corresponding to R encryption rounds (R will depend on the evaluation results of the safety of each algorithm). The general design will be depicted in part a, while part b and part c will demonstrate the individual content of each algorithm.

a. The general algorithm and the scheme of TMN

The algorithm:

1. For $j = 1$ to $R - 1$ do $\{(A, B) \rightarrow Crypt^{(e)}(A, B, Q_j^{'}); (A, B) \rightarrow (B, A)\}$
2. $\{(A, B) \rightarrow Crypt^{(e)}(A, B, Q_R^{'})\}$
3. $\{(A, B) \rightarrow (A \rightarrow Q_{R+1}^{'}, B \rightarrow Q_{R+1}^{'})\}$.

In which: R is the number of rounds of the algorithm; $R+1$ is relative to the final transformation (FT) (the final transformation is conducted by running the XOR on half of left and right data blocks with the corresponding round key born); $e \in 0, 1$ is the parameter which is defined as: $(e = 0)$ is encrypting and $(e = 1)$ is decrypting; $Crypt^{(e)}$ is the function which describes the transformation of the base encryption round of each algorithm. This function is dissimilar to each algorithm and is depicted in the scheme of each algorithm.

The scheme of the algorithm. With the way of designing the components of each algorithm and the involution of each component, these algorithms allow the encryption and decryption process to use the same scheme. The description of the 2 algorithms is indicated on Fig. 1.

Figure 1b and c describe the design of one base encryption round (the transforming function $Crypt^{(e)}$) of the algorithms **TMN64** and **TMN128** respectively.

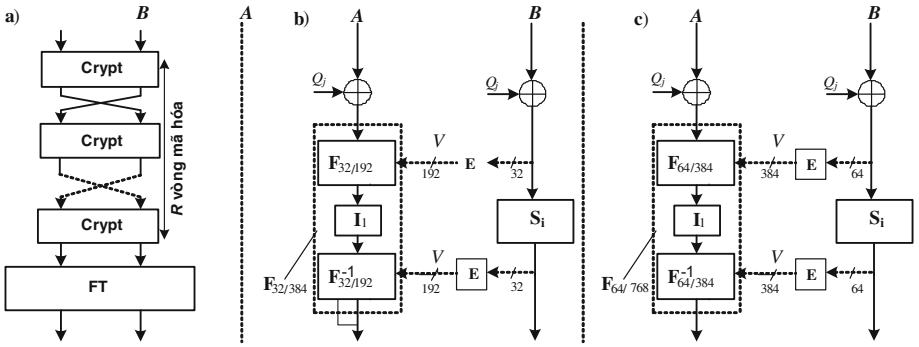


Fig. 1. The scheme of the proposed algorithms; (a) General look; (b) The base encryption round of **TMN64**; (c) The base encryption round of **TMN128**.

b. The specific point in the design of the algorithm TMN64

The left branch of the algorithm **TMN64** use 2 CPSNs: $F_{32/192}, F^{-1}_{32/192}$. The $F_{32/192}$ is designed based on 2 other CSPNs: $F_{32/96}$ (including 4 blocks of $F_{8/24}$), $F^{-1}_{32/96}$ (including 4 blocks of $F^{-1}_{8/24}$), whereas the $F^{-1}_{32/192}$ is the connection between 2 CSPNs: $F^{-1}_{32/96}$ (including 4 blocks of $F^{-1}_{8/24}$, 4 blocks of $F_{8/24}$).

The right branch of the algorithm **TMN64** apply CSPN S_i, which is designed as described on Fig. 2. The expansion block E is designed as below: $E(X) = (X, X^{<<6}, X^{<<12}, X^{<<18}, X^{<<24}, X^{<<30})$, in which $^{<<b}$ is the left shift rotating b bits.

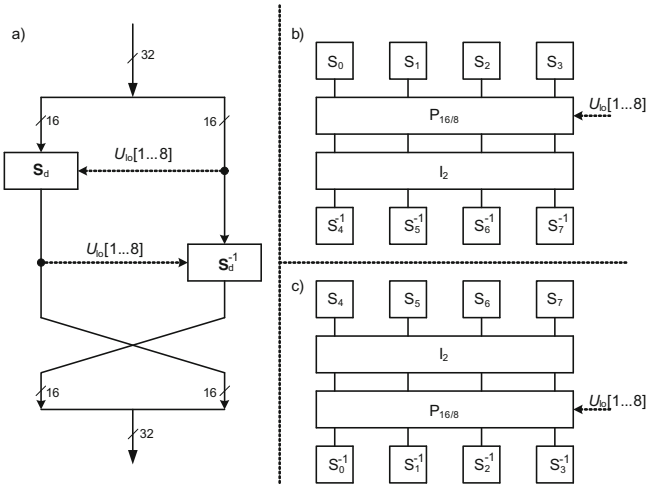


Fig. 2. CSPN $\mathbf{S}_i$ (a); block $\mathbf{S}_d$ (b); block $\mathbf{S}_d^{-1}$(c)

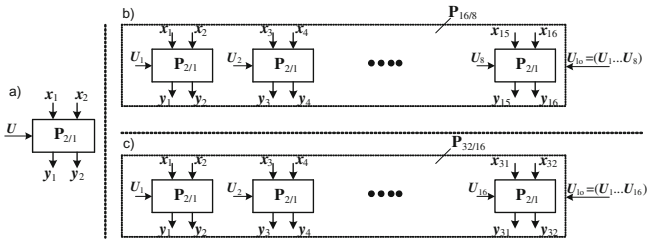


Fig. 3. The controllable block $P_{16/8}$

c. The specific point in the design of the algorithm TMN128
The left branch of the algorithm TMN128 use 2 CSPNs $F_{64/384}$ and $F^{-1}_{64/384}$. The $F_{32/384}$ is designed based on 2 other CSPNs: $F_{64/192}$ (including 8 blocks of $F_{8/24}$) and $F^{-1}_{64/192}$ (including 8 blocks of $F^{-1}_{8/24}$), while the $F^{-1}_{64/384}$ is the connection between 2 CSPNs $F^{-1}_{64/192}$ (including 8 blocks of $F^{-1}_{8/24}$, 8 blocks of $F_{8/24}$).

The right branch of the algorithm **TMN128** use CSPN S_i. The expansion block E is designed as below: $E(X) = (X, X^{<<12}, X^{<<24}, X^{<<36}, X^{<<48}, X^{<<60})$, in which $^{<<b}$ is the left shift rotating b bits. The boxes $S_{4\times4}$ is used as: the proposed algorithms will use the boxes $S_{4\times4}$ (which have been proved to be safe) of the algorithm Serpent-1 [8].

2.2 Designing the Key Schedule

The method of expanding the secret key is the simple yet effective solution to prevent the weaknesses of the key, according to the theory of the cipher blocks [2]. However, conducting the expansion of the key requires more resources (when conducting with devices) or needs to premeditate all the round keys before starting to encrypt, which will reduce the encryption speed when having to change the secret key frequently. To surmount these issues, we need to propose the technique that creates the on-the-fly round key. With this solution, the scheme of the encryption algorithm will have a blockwhich expands the key, in order to create the next round key in the time of implementing the encryption (decryption) round.

The general scheme of the encryption algorithms when using the on-the-fly key schedule is described on Fig. 4.

Detailed description ***about the procedure of creating the round key*** (for the algorithm **TMN**):

1. At the front, the first encryption round applies the switching operator $T^{(e)}$, implements the transformation on 32 bits (or 64 bits) correspondingly to the algorithms, locks (k_1, k_2) to create the round key $Q_1 = (X, Y)$, in which the secret key is in the form: $K = (k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8)$. Notice that, in the encryption mode $(e = 0), T^{(e)}(X, Y) = (X, Y)$ and in the decryption mode $(e = 1), T^{(e)}(X, Y) = (Y, X)$.
   Therefore, in the encryption mode: $Q = (k_1, k_2)$ and in the decryption mode: $Q1 = (k_2, k_1)$.

2. Round $(R/2+1)$ does not use the switching operator $T^{(e)}$.
3. The remaining rounds $(1...R/2, R/2+2...R+1)$ will use the switching operator $T^{(e)}$.
4. After round $(R/2+1)$ uses the switching operator $T^{(1)}$ (the switching operator remains consistent and unchanged in both the encryption and decryption). The input of $T^{(1)}$ receives the output value of block $Q^{(e)}_exp$ (at $z_{h/2}$), whereas the output of $T^{(1)}$ lays on the input of block Q_exp (at $z_{h/2+1}$). Notice that, block Q_exp remains unchanged when altering from the encryption to the decryption, which means it does not depend on the value $e=1$ or $e=0$.

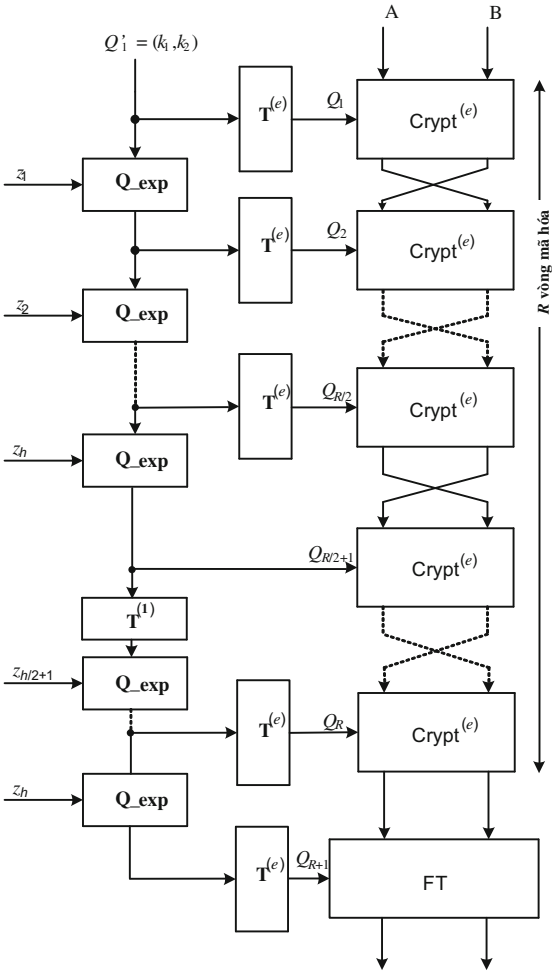


Fig. 4. The general scheme of the proposed algorithms when using the on-the-fly round key schedule

5. In the encryption round $Crypt^{(e)}$, data block A and data block B will be added with the round key $Q_j = (X, Y)$. Which means: with data block A : $A \leftarrow A \oplus (X, Y)$, with data block $B : B \leftarrow B \oplus (X, Y)$.
6. The key values z_h (not depending on how it is created) while encrypting/decrypting will be applied in the opposite order.

The success of the plan which operates the on-the-fly key schedule to create the expansion key is attributed to the usage of the switching operator. This plan is presented on Fig. 5. Notice that, with respect to the algorithm **TMN64**, $m = 32$ and $n = 16$, with respect to the algorithm **TMN128**, $m = 64$ and $n = 32$.

The procedures are repeated to receive the expansion key Q_exp which contains the blocks S_d and S_d^{-1}, while the right branch applies the encryption rounds (Fig. 4). The blocks S_d and S_d^{-1} are designed using the substitution boxes S_i, S_i^{-1} whose sizes are 4×4, and $CEP_{2/1}$. The blocks S_d and S_d^{-1} are depicted on Fig. 2b, c, respectively to the algorithms **TMN64** and **TMN 128**.

The procedure Q_exp does not have the involution, but is a switching operator, which means $Q_exp = (Q_exp)^{-1}$.

The switching feature of Q_exp is gained by using the transposition block $T^{(e)}$ on the scheme of Fig. 5 (it actually is the transposition block or correspondingly to the algorithms **TMN64** and **TMN128**)

The switching element $T^{(e)}$ implements in accordance with the formula: $T^{(e)}(X, Y) = (X, Y)$ if $e = 0$ and $T^{(e)}(X, Y) = (Y, X)$ if $e = 1$. In here, $X, Y \in \{0, 1\}^{16}$ or $X, Y \in \{0, 1\}^{32} e = 0 - encrypting, e = 1 - decrypting$ (relatively to the algorithms).

*In case we use the secret key with 128 bits (the algorithm **TMN64**)*: the procedure which creates the intricate round key uses the secret key in the form: $K = (k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8)$.

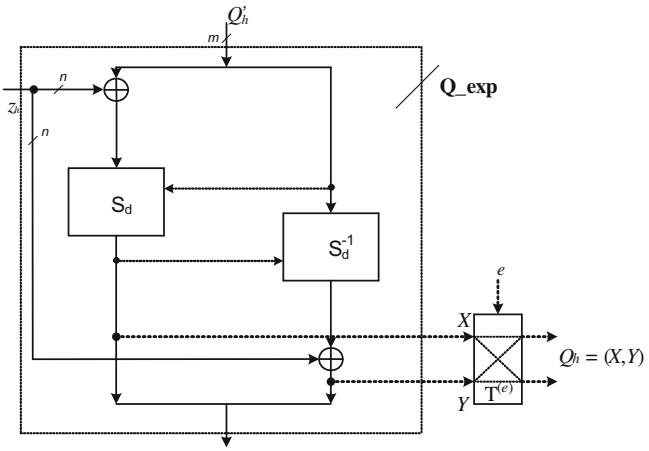


Fig. 5. The procedure Q_exp to expand the secret key to the on-the-fly round key

Table 1. The key schedule in TMN64 (with R = 10)

The rounds	1	2	3	4	5	6	7	8	9	10	11
Value h	1	2	3	4	5	6	7	8	9	10	-
Encrypting z_h	k_7	k_3	k_5	k_1	k_6	k_6	k_1	k_5	k_3	k_7	-
Decrypting z_h	k_7	k_3	k_5	k_1	k_6	k_6	k_1	k_5	k_3	k_7	-

Table 2. The key schedule in **TMN128** (with $R = 12$)

Round j	1	2	3	4	5	6	7	8	9	10	11	12	13
Value h	1	2	3	4	5	6	7	8	9	10	11	12	13
Encrypting z_h	k_3	k_5	k_1	k_6	k_4	k_2	k_2	k_4	k_6	k_1	k_5	k_3	-
Decrypting z_h	k_3	k_5	k_1	k_6	k_4	k_2	k_2	k_4	k_6	k_1	k_5	k_3	-

The values $z_1, z_2, ..., z_h$, used in the procedure which creates the intricate round key are shown in the Table 1 (in here the subkeys $(k_1, k_2, ..., k_8)$ have the size of 16 bits.)

*In case we use the secret key with 256 bits (the algorithm **TMN128**)*: the procedure which creates the intricate round key uses the secret key in the form: $K = (k_1, k_2, k_3, k_4, k_5, k_6, k_7, k_8)$. The values $z_1, z_2, ..., z_h$, used in the procedure which creates the intricate round key are shown in the Table 2 (in here the subkeys $(k_1, k_2, ..., k_8)$ have the size of 32 bits).

3 Analyzing the Security of the Proposed Algorithms

3.1 Estimating the Statistical Standards

To estimate the statistical standards of the algorithm based on CE $F_{2/2}$, this part will implement the experiments in accordance with the method described in [9]. The statistically analyzing results of **TMN64**, **TMN128** are operated similarly to the experiments in the algorithms attending the final round of AES [9]. The results show that after 4 rounds of encryption they can satisfy the criteria of NESSIE.

3.2 Assessing the Safety to the Differential Cryptanalysis

The results of evaluating the differential cryptanalysis of the algorithm **TMN64** have shown that, the greatest differential feature is corresponding to the two-round differential value with one active bit. For that reason, the feature has the probability $P(2) \leq 2^{-36}$. It is specifically calculated as below. One active bit (in the left branch) Δ^{A_1} will go through the first round with the probability: $p_1(\Delta^{A_1} \rightarrow \Delta^{A_1}) = (P(ijk))^{12} = (P(110))^{12} = 0.75^{12} \approx 2^{-5}$. It will be constituted at the output of the value Δ^{A_1} and then taken to the right branch of the

second round. The right branch of the first round will carry out the transformation through CSPN S_i with the differential $(\Delta^{B_0} \rightarrow \Delta^{B_0})$ and the probability is $p_2 = 1$. One active bit Δ^{B_1} going through the second round (through S_i) will create Δ^{B_1} with the probability: $p_4(\Delta^{B_1} \rightarrow \Delta^{B_1}) = P(\Delta^{A_1} \underline{S_i} \Delta^{A_1}) = 2^{-7}$. The blocks $F_{32/192}, F^{-1}_{32/192}$ (at the second round) will be controlled by 3 active bits (created through the expansion block **E**). Consequently, the probability of the differential $\Delta^{A_0} \rightarrow \Delta^{A_0}$ through the block $F_{32/192}$ will be: $p_3(\Delta^{A_0} \rightarrow \Delta^{A_0}) = (P(001))^6 = 2^{-12}$. Similarly, the probability of the differential $\Delta^{A_0} \rightarrow \Delta^{A_0}$ through the block $F^{-1}_{32/96}$ will be: $p5(\Delta^{A_0} \rightarrow \Delta^{A_0}) = (P(001))^6 = 2^{-12}$. Hence, the probability of the differential feature after 2 rounds of **TMN64** will be: $P(2) = p_1.p_2.p_3.p_4.p_5 \leq 2^{-36}$. *Therefore, to ensure the securyty before the attacks of the differential, the algorithm* ***TMN64*** *needs to use 10 encryption rounds.*

With respect to the algorithm **TMN128**: the proofs about the differential cryptanalysis will be operated similarly. There is only one discrepancy: the probability of the differential going through the block Si will be smaller than 2–8. Thus, the probability of the differential feature after 2 rounds of **TMN128** will be: $P(2) = p_1.p_2.p_3.p_4.p_5 \leq 2^{-37}$. *Therefore, to ensure the safety before the attacks of the differential, the algorithm* ***TMN128*** *needs to use 12 encryption rounds.*

4 Estimate Performance of the Proposed Algorithm on FPGA

The models which evaluate the integrated efficiency of the algorithms TMN are conducted based on the models in [2,4]. These algorithms are implemented on FPGA according to the iterative structure and the pipeline structure (in the ECB mode). The received results are in comparison with the algorithms AES [9],

Table 3. FPGA Synthesis Results of **TMN** and Comparisons. (1): Order. (2): Algorithm. (3): Data block (bit). (4): Number of rounds of algorithm. (5): Resources (CLBs). (6): Frequency (MHz). (7): Throughput (Mbps). (8): $\frac{Mbps}{CLBs}$. (9): $\frac{Mbps}{CLBs.Ghz}$

(1)	(2)	(3)	(4)	(5)	(6)	(7)	Efficiency	
							(8)	(9)
I	The iterative structure							
1	**TMN64**	64	10	82	136	871	10.61	78.05
2	**TMN128**	128	12	132	133	1419	10.75	80.81
3	**AES**	128	10	598	225	2880	4.82	21.4
II	The pipeline structure							
1	**TMN64**	64	10	368	150	992	26.96	173.91
2	**TMN128**	128	12	805	153	1996	26.2	167.98
3	**AES**	128	10	3572	58	7424	2.1	35.8

and have proved that, the proposed algorithms have higher operating speed and integrated efficiency than those in the final round of AES. Simultaneously, the received results of assessing the performance have also proved that the algorithms **TMN** are extremely appropriate for the integration on FPGA and suitable for being used in WMSs (Table 3).

5 Conclusion

The paper has attained some main results as: constructing the high-speed block ciphers **TMN** based on the CE $F_{2/2}$ and proposing a plan to build up the schedule which creates the on-the-fly round key. With this schedule, the proposed algorithms still ensure high efficiency when regularly change of the session key and counteracting the greatest weakness of the algorithms based on CSPN, on account of using the schedule which creates the simple key. The results that evaluate the safety of the algorithms by the statistical standard NESSIE and the differential cryptanalysis have shown that these are the algorithms safe enough in accordance with the popular assessing criterias in the world. Concurrently, the paper has also conducted the assessment on the integrated efficiency of the proposed algorithms on FPGA. The received results have proved that the proposed algorithms are effective enough when deploying the operation on the hardware of FPGA and can be applied in Wireless Communication Networks.

References

1. Kong, J., Gerla, M.: Adaptive security support for real-time transmissions in ad hoc networks. http://netlab.cs.ucla.edu/wiki/files/MEDHOCNET-jkong.pdf
2. Moldovyan, N., Moldovyan, A.: Data-driven Block Ciphers for Fast Telecommunication Systems. Auerbach Publications, New York (2008)
3. Moldovyan, N., Moldovyan, A.: Innovative Cryptography. Charles River Media, Boston (2007)
4. Minh, N.H., Bac, D.T., Duy, H.N.: New SDDO-based block cipher for wireless sensor network security. Intl. J. Comput. Sci. Netw. Secur. **10**(3), 54–60 (2010)
5. Bodrov, A.V., Moldovyan, A.A., Moldovyanu, P.A.: DDP-based cipers: differential analysis of SPECTR-H64. Comput. Sci. J. Moldova **13**(3), 268–291 (2005). 39
6. Kang, J., Jeong, K., Yeo, S.-S., Lee, C.: Related-key attack on the MD-64 block cipher suitable for pervasive computing environments. In: 26th IEEE International Conference on Advanced Information Networking and Applications Workshops (WAINA), pp. 726–731. IEEE Xplore (2012)
7. Kang, J., Jeong, K., Lee, C., Hong, S.: Distinguishing attack on SDDO-based block cipher BMD-128. In: Jeong, Y.-S., Park, Y.-H., Hsu, C.-H.R., Park, J.J.J.H. (eds.) Ubiquitous Information Technologies and Applications (CUTE 2013). LNEE, vol. 280, pp. 595–602. Springer, Heidelberg (2014). doi:10.1007/978-3-642-41671-2_76
8. Anderson, R., Biham, E., Knudsen, L.: Serpent: a proposal for the advanced encryption standard. http://cryptosoft.net/docs/Serpent.pdf
9. NESSIE. New European Schemes for Signatures, Integrity, and Encryption. https://www.cosic.esat.kuleuven.ac.be/nessie/
10. Kang, J., Jeong, K., Hong, S., Lee, C.: Related-key amplified boomerang attacks on KT-64 and MD-64 suitable for wireless sensor networks. Sensor Lett. **11**(9), 1765–1770 (2013)

Numerical Method for Solving a Strongly Mixed Boundary Value Problem in an Unbounded Domain

Dang Quang A[1] and Dinh Hung Tran[2(✉)]

[1] Centre for Informatics and Computing, Vietnam Academy of Science and Technology (VAST), 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam
dangquanga@cic.vast.vn
[2] College of Education, Thai Nguyen University, Thai Nguyen City, Vietnam
trandinhhungvn@gmail.com

Abstract. In this paper we consider a strongly mixed elliptic boundary value problem in an unbounded domain, where there is one point of transmission of boundary conditions on an unbounded boundary. By a domain decomposition technique the problem is reduced to two problems, one with weakly mixed boundary conditions in a bounded subdomain and another with the Dirichlet boundary conditions in the remaining unbounded subdomain. For the problem in the bounded subdomain the finite difference method on a uniform grid is used. Meanwhile the problem in the unbounded subdomain is discretized on a non-uniform grid and for solving the resulting system of grid equations we use the method of infinite systems, developed by ourselves recently. Numerical experiments, where the grid with the monotonically increasing grid sizes on the unbounded subdomain is used, show the efficiency of the proposed method.

Keywords: Strongly mixed boundary value problems · Unbounded domain · Domain decomposition method · Finite difference · Infinite system

1 Introduction

In this paper we consider a strongly mixed elliptic boundary value problem in a semistrip, where there is one point of transmission of boundary conditions on an unbounded boundary:

$$\begin{aligned}
&Lu = \gamma_1 \frac{\partial^2 u}{\partial x^2} + \gamma_2 \frac{\partial^2 u}{\partial y^2} - bu(x,y) = f(x,y), \quad x > 0,\ 0 < y < 1,\\
&u(0,y) = \varphi_1(y),\ 0 \le y \le 1,\\
&\frac{\partial u}{\partial y}(x,0) = \psi(x),\ 0 \le x \le L,\\
&u(x,0) = \varphi_2(x),\ x \ge L,\\
&u(x,1) = \varphi_3(x),\ x \ge 0,\\
&u(x,y) \to 0, \quad x \to +\infty,
\end{aligned} \tag{1}$$

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_44

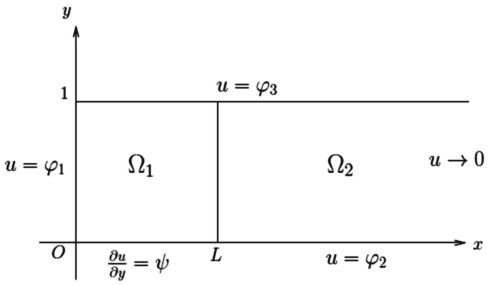


Fig. 1. Boundary conditions and subdomains

under the usual assumptions that the functions in (1) are continuous and

$$\gamma_1, \gamma_2 > 0, \quad b \geq 0, \quad f(x, y) \to 0, \ \varphi_i(x) \to 0, \quad x \to +\infty, \ (i = 2, 3).$$

This problem is a model of the stationary two-dimensional problem of mass propagation in a semistrip when the flow is in the x-direction. It ought to be noted that the assumptions of compact support of the source function and the boundary conditions are not required.

It should be emphasized that the problem (1) belongs to the class of problems with possible boundary singularities at the points of smooth parts of boundary, where there happen changes of types of boundary conditions. We call them strictly mixed boundary value problems for distinguishing from those, in which the changes of types of boundary conditions occur at angle points. Many problems of this class are usually reduced to dual series or integral equations (see e.g., [6,7,16], which later in their turn are transformed to Fredholm integral equation of the second kind for numerical solution.

On this occassion we remark that for solving Laplacian and elliptic problems with boundary singularities in bounded domains several effective methods have been developed, among them the singular function boundary integral method attracts great attention (see, e.g., [8,9,11,12]. In this method the solution is approximated by the leading terms of the local asymptotic solution expansion near the singular point.

In the present work, using the idea of the domain decomposition method, developed recently in [5] for solving strongly mixed boundary value problems, we reduce the problem (1) to two problems, one with weakly mixed boundary conditions in a bounded subdomain and another with the Dirichlet boundary conditions in the remaining unbounded subdomain.

For the problem in the bounded subdomain we use the difference method of second order approximation, and use the method of complete reduction for solving the resulting difference equations. Meanwhile, for the problem in the unbounded subdomain after discretization we obtain an infinite system of three-point vector equations. Applying the idea of Polozhii in the method of summary representations [13] we transform this system to systems of three-point scalar

equations and use the truncation method for obtaining approximate solutions with a given accuracy. This technique was used in our recent work [2]. But in this work the use of uniform grids leads to a large computational account, therefore, in its Conclusion we have a remark of the possibility of using monotonically increasing grid sizes for improving the efficiency of the method of infinite systems. In the present paper we realize this idea. Numerical experiments, where the grid with the monotonically increasing grid sizes on the unbounded subdomain is used, show the efficiency of the proposed method.

2 Iterative Method on Continuous Level

Divide Ω into two subdomains Ω_1, Ω_2 by the smooth boundary Γ, so that $\Omega = \Omega_1 \bigcup \Omega_2$, $\Omega_1 \bigcap \Omega_2 = \emptyset$ (see Fig. 1). Denote by u_i the solutions on the subdomains Ω_i $(i = 1, 2)$.

Consider the following iterative method for finding u_1 and u_2 in the subdomains based on the updating of $g = \frac{\partial u_1}{\partial x}$ on the interface Γ:
(1) Given $g^{(0)} \in L^2(\Gamma)$, for example $g^{(0)} = 0, 0 \le y \le 1$.
(2) Knowing $g^{(k)}$ on Γ $(k = 0, 1, 2, ...)$ solve consecutively two problems

$$\begin{aligned}
&Lu_1^{(k)} \equiv \gamma_1 \frac{\partial^2 u_1^{(k)}}{\partial x^2} + \gamma_2 \frac{\partial^2 u_1^{(k)}}{\partial y^2} - b u_1^{(k)}(x,y) = f(x,y),\ 0 < x < L,\ 0 < y < 1,\\
&u_1^{(k)}(0,y) = \varphi_1(y),\ 0 \le y \le 1,\\
&\frac{\partial u_1^{(k)}}{\partial y}(x,0) = \psi(x),\ 0 \le x \le L,\\
&u_1^{(k)}(x,1) = \varphi_3(x),\ 0 \le x \le L,\\
&\frac{\partial u_1^{(k)}}{\partial x}(L,y) = g^{(k)}(y),\ 0 \le y \le 1.
\end{aligned} \tag{2}$$

$$\begin{aligned}
&Lu_2^{(k)} \equiv \gamma_1 \frac{\partial^2 u_2^{(k)}}{\partial x^2} + \gamma_2 \frac{\partial^2 u_2^{(k)}}{\partial y^2} - b u_2^{(k)}(x,y) = f(x,y),\ x > L,\ 0 < y < 1,\\
&u_2^{(k)}(x,0) = \varphi_2(x),\ x \ge L,\\
&u_2^{(k)}(x,1) = \varphi_3(x),\ x \ge L,\\
&u_2^{(k)}(L,y) = u_1^{(k)}(L,y),\ 0 \le y \le 1,\\
&u(x,y) \to 0,\ x \to +\infty.
\end{aligned} \tag{3}$$

(3) Compute the new approximation $g^{(k+1)}$

$$g^{(k+1)} = (1-\tau) g^{(k)} + \tau \frac{\partial u_2^{(k)}}{\partial x}(L,y),\ 0 \le y \le 1, \tag{4}$$

where τ is an iterative parameter to be chosen.

Remark 1. The above iterative method is an extension of the domain decomposition method developed in [5] for the Poisson equation and in [3,4] for the biharmonic equation in bounded domains to unbounded ones. There, the method is proved to be convergent for any initial approximation.

3 Numerical Realization of the Iterative Method

We see that at each iteration of the method described above it is needed to solve two boundary value problems, one in a bounded subdomain, and another in an unbounded subdomain. For solving them we shall use the difference method [14]. To do this we construct the uniform grid $\overline{\omega}_h^1$ in the bounded subdomain Ω_1 and the non-equidistant grid $\overline{\omega}_h^2$ in $x-$direction in the unbounded subdomain Ω_2 as follows

$$\overline{\omega}_h^1 = \{(x_i, y_j),\ x_i = ih_1,\ y_j = jh_2,\ i = 0, 1, ..., N;\ j = 0, 1, ..., M\},$$

$$\overline{\omega}_h^2 = \left\{(\widehat{x}_i, y_j),\ \widehat{x}_{i+1} = \widehat{x}_i + \widehat{h}_{i+1},\ y_j = jh_2,\ i = 0, 1, ...;\ j = 0, 1, ..., M\right\},$$

where $N = 2^n,\ h_1 = L/N,\ h_2 = 1/M,\ \widehat{x}_0 = L.$

Here, we use the non-equidistant grid in the unbounded subdomain Ω_2 following a remark in Conclusion of [2].

For solving the weakly mixed BVP (2) on the grid $\overline{\omega}_h^1$ we use a standard function in the program package TK2004 constructed by Vu in [17], where the method of complete reduction [15] is used for solving grid equations. The solution of the difference scheme of (2) is denoted by $v_{i,j}^{1(k)}$, i.e., $v_{i,j}^{1(k)} \approx u_1^{(k)}(x_i, y_j)$.

We also denote by $v_{i,j}^{2(k)}$ the approximation of the values $u_2^{(k)}(\widehat{x}_i, y_j)$ on the grids $\overline{\omega}_h^2$, and set

$$f_{ij} = f(\widehat{x}_i, y_j),\ (\widehat{x}_i, y_j) \in \overline{\omega}_h^2,\ g_j^{(k)} = g^{(k)}(y_j),\ j = 0, 1, ..., M.$$

For solving the problem (3), we replace it by the difference scheme

$$\begin{aligned} &L_h v_2^{(k)} \equiv \gamma_1 v_{\overline{x}\widehat{x}}^{2(k)} + \gamma_2 v_{\overline{y}y}^{2(k)} - bv^{2(k)} = f_{ij},\ i = 1, 2, ...;\ j = 1, 2, ..., M-1 \\ &v_{i,0}^{2(k)} = \varphi_2(\widehat{x}_i),\ v_{i,M}^{2(k)} = \varphi_3(\widehat{x}_i),\ v_{0,j}^{2(k)} = v_{N,j}^{1(k)},\ i = 0, 1, ...;\ j = 0, 1, ..., M \\ &v_{i,j}^{2(k)} \to 0,\ i \to +\infty;\ j = 0, 1, ..., M, \end{aligned} \tag{5}$$

where

$$v_{\overline{x}\widehat{x}}^{2(k)} = v_{\overline{x}\widehat{x}}^{2(k)}(\widehat{x}_i, y_j) = \frac{1}{\hbar_i}\left(\frac{v_{i+1,j}^{2(k)} - v_{i,j}^{2(k)}}{\widehat{h}_{i+1}} - \frac{v_{i,j}^{2(k)} - v_{i-1,j}^{2(k)}}{\widehat{h}_i}\right)$$

$$\hbar_i = \frac{\widehat{h}_i + \widehat{h}_{i+1}}{2},\ i = 1, 2, ...$$

Now, denote

$$V_0^{(k)} = \begin{pmatrix} v^{1(k)}(N, y_1) \\ v^{1(k)}(N, y_2) \\ \dots \\ v^{1(k)}(N, y_{M-1}) \end{pmatrix},$$

$$V_i^{(k)} = \begin{pmatrix} v_{i,1}^{2(k)} \\ v_{i,2}^{2(k)} \\ \dots \\ v_{i,M-2}^{2(k)} \\ v_{i,M-1}^{2(k)} \end{pmatrix}, \quad \overline{F}_i = \begin{pmatrix} f_{i,1} - \frac{\gamma_2}{h_2^2}\varphi_2(x_i) \\ f_{i,2} \\ \dots \\ f_{i,M-2} \\ f_{i,M-1} - \frac{\gamma_2}{h_2^2}\varphi_3(x_i) \end{pmatrix}, \quad i = 1, 2, \dots.$$

The Eq. (5) can be written in the form of three-point vector difference equations

$$A_i V_{i-1}^{(k)} + \frac{\gamma_2}{h_2^2} T V_i^{(k)} + B_i V_{i+1}^{(k)} - (A_i + B_i + b + 2\frac{\gamma_2}{h_2^2}) V_i^{(k)} = \overline{F}_i \quad i = 1, 2, \dots \quad (6)$$

where

$$A_i = \frac{\gamma_1}{\hbar_i \widehat{h}_i}, \quad B_i = \frac{\gamma_1}{\hbar_i \widehat{h}_{i+1}}, \ i = 1, 2, \dots \quad (7)$$

$V_i^{(k)} \to 0$ when $i \to +\infty$ and T is the matrix of order $M-1$

$$T = \begin{pmatrix} 0 & 1 & 0 & 0 & \dots & 0 & 0 & 0 \\ 1 & 0 & 1 & 0 & \dots & 0 & 0 & 0 \\ 0 & 1 & 0 & 1 & \dots & 0 & 0 & 0 \\ . & . & . & . & \dots & . & . & . \\ . & . & . & . & \dots & . & . & . \\ 0 & 0 & 0 & 0 & \dots & 1 & 0 & 1 \\ 0 & 0 & 0 & 0 & \dots & 0 & 1 & 0 \end{pmatrix}.$$

Next, we shall use the idea of Polozhii in the method of summary representations [13] to transform the infinite system of three-point vector equations (6) to infinite systems of three-point scalar equations. For this purpose let us introduce the notations

$$S = (s_{ij})_1^{M-1}, \quad s_{ij} = \sqrt{\frac{2}{M}} \sin \frac{ij\pi}{M}, \quad i, j = 1, 2, \dots, M-1,$$

$$\Lambda = [\lambda_1, \lambda_2, \dots, \lambda_{M-1}], \quad \lambda_j = 2\cos\frac{j\pi}{M}, \quad j = 1, 2, \dots, M-1.$$

We have $S^T = S$, $S^2 = E$ and $T = S^{-1}\Lambda S$.

Multiply (6) by S and put $W_i^{(k)} = (w_{i,j}^{(k)}) = SV_i^{(k)}$, $G_i = (g_{i,j}) = S\overline{F}_i$, $i = 0, 1, 2, \dots$, $j = 1, 2, \dots, M-1$. Then for every fixed index j we have the system of customary three-point difference equations

$$\begin{aligned} A_i w_{i-1,j}^{(k)} - C_{i,j} w_{i,j}^{(k)} + B_i w_{i+1,j}^{(k)} &= -F_{i,j}, \quad i = 1, 2, \dots \\ w_{0,j}^{(k)} = \mu_{0,j}^{(k)}, \ w_{i,j}^{(k)} &\to 0, \quad i \to \infty, \end{aligned} \quad (8)$$

where the numbers A_i, B_i are defined by (7), $\mu_{0,j}^{(k)} = \sum_{l=1}^{M-1} s_{j,l} v_{0,l}^{(k)}$, $F_{i,j} = -g_{i,j}$ and

$$C_{i,j} = A_i + B_i + b + \frac{4\gamma_2}{h_2^2} \sin^2 \frac{j\pi}{2M} > 0.$$

Further, in order to treat the system (8) we shall use the method of infinite system of equations in [10], which was developed by ourselves for solving some one-dimensional problems in [1]. For this purpose we set

$$\begin{aligned} p_{0,j}^{(k)} = q_{0,j}^{(k)} = 0,\ r_{0,j}^{(k)} = \mu_0^{(k)}, \\ p_{i,j}^{(k)} = \frac{A_i}{C_{i,j}},\ q_{i,j}^{(k)} = \frac{B_i}{C_{i,j}},\ r_{i,j}^{(k)} = \frac{F_{i,j}}{C_{i,j}},\ i = 1, 2, ... \end{aligned} \tag{9}$$

and rewrite the system (8) in the canonical form of infinite system as follows

$$\begin{aligned} w_{i,j}^{(k)} &= p_{i,j}^{(k)} w_{i-1,j}^{(k)} + q_{i,j}^{(k)} w_{i+1,j}^{(k)} + r_{i,j}^{(k)},\ i = 0, 1, 2, ... \\ w_{i,j}^{(k)} &\to 0,\ i \to \infty. \end{aligned} \tag{10}$$

The system (10) is regular and the solution of this infinite system can be found by the truncation method. Following the progonka method (or Thomas algorithm) which is a special form of the Gauss elimination [14] for tridiagonal system of equations we shall seek the solution of (10) in the form

$$w_{i,j}^{(k)} = \alpha_{i+1,j}^{(k)} w_{i+1,j}^{(k)} + \beta_{i+1,j}^{(k)},\ i = 0, 1, ..., \tag{11}$$

where coefficients are calculated by the formulas

$$\begin{aligned} \alpha_{1,j}^{(k)} = 0,\ \beta_{1,j}^{(k)} = \mu_0^{(k)}, \\ \alpha_{i+1,j}^{(k)} = \frac{q_{i,j}^{(k)}}{1 - p_{i,j}^{(k)} \alpha_{i,j}^{(k)}},\ \beta_{i+1,j}^{(k)} = \frac{r_{i,j}^{(k)} + p_{i,j}^{(k)} \beta_{i,j}^{(k)}}{1 - p_{i,j}^{(k)} \alpha_{i,j}^{(k)}},\ i = 1, 2, ... \end{aligned} \tag{12}$$

There holds the following

Theorem 1. *If for a given $\varepsilon > 0$, there exists a natural number $N^{(k)}$ such that*

$$\frac{|\beta_{i,j}^{(k)}|}{1 - \alpha_{i,j}^{(k)}} \le \varepsilon,\ \forall i \ge N^{(k)} + 1;\ j = 1, 2, ..., M - 1$$

then for the solution of the truncated system

$$\begin{aligned} &\bar{w}_{i,j}^{(k)} = p_{i,j}^{(k)} \bar{w}_{i-1,j}^{(k)} + q_{i,j}^{(k)} \bar{w}_{i+1,j}^{(k)} + r_{i,j}^{(k)},\ i = 0, 1, 2, ..., N^{(k)},\ j = 1, 2, ..., M - 1, \\ &\bar{w}_{i,j}^{(k)} = 0,\ i \ge N^{(k)} + 1,\ j = 1, 2, ..., M - 1 \end{aligned}$$

there holds the estimate

$$\sup_{i,j} |w_{i,j}^{(k)} - \bar{w}_{i,j}^{(k)}| \le \varepsilon, \tag{13}$$

where $w_{i,j}^{(k)}$ is the solution of the infinite system (10).

The above theorem follows from Theorem 3.2 in [1].

Now denote $\bar{V}_i^{(k)} = (\bar{v}_{i,j}^{(k)})_{j=1}^{M-1}$, $\bar{W}_i^{(k)} = (\bar{w}_{i,j}^{(k)})_{j=1}^{M-1}$, $i = 0, 1, 2, ...$ and set $\bar{V}_i^{(k)} = S\bar{W}_i^{(k)}$. Then, as in [2], we have the following result.

Theorem 2. *There holds the estimation*

$$\sup_{i,j} |v_{i,j}^{2(k)} - \bar{v}_{i,j}^{(k)}| < \sqrt{M-1}\,\varepsilon. \tag{14}$$

It should be remarked that the above theorem serves as a criterion for automatic truncating the infinite system in the process of computation of the sweep coefficients (12) for guaranteing a given accuracy of the obtained solution.

4 Numerical Examples

Below we present the results of numerical experiments to illustrate the effectiveness of the proposed iterative method (2)–(4) on some examples of the problem (1) with known exact solutions. All the corresponding right hand side functions tend to zero at infinity, $\gamma_1 = \gamma_2 = b = 1$, $L = 1$. The experiments are performed on uniform grids in the bounded subdomain with grid sizes h_1, h_2 and on the non-equidistant grids in $x-$direction in the unbounded subdomain with increasing grid sizes, $\widehat{h}_1 = h_1$, $\widehat{h}_{i+1} = 1.2\widehat{h}_i$, $i = 1, 2, ...$ In the results reported $N^{(k)}$ is the size of the systems that is automatically truncated with the given accuracy ε in Theorem 1, $\text{error}_1 = \max\limits_{i,j} |u_1(x_i, y_j) - v_{i,j}^{1(k)}|$, $\text{error}_2 = \max\limits_{i,j} |u_2(x_i, y_j) - \bar{v}_{i,j}^{2(k)}|$ are the errors of the obtained approximate solution compared with the exact solution in bounded and unbounded subdomains, Ω_1 and Ω_2 respectively.

Besides, for demonstrating the working capacity of the method we consider one example, where the exact solutions are unknown.

4.1 Example 1

For simplicity we take

$$u = \frac{e^{-y}}{x+1} + ye^{-x}.$$

The results of convergence for $N = 2^4 = 16, h_1 = h_2 = 1/N = 0.0625$, $\varepsilon = 0.01$ with a number of different values of the iterative parameter τ and k are given in the Table 1. From this table we see that the good approximate solution is obtained when τ varies from 0.5 to 0.7.

4.2 Example 2

Now we take

$$\varphi_1(y) = y^2, \psi(x) = 0, \varphi_2(x) = e^{-2x}, \ \varphi_3(x) = \frac{1}{x+1}, \ f(x,y) = \frac{1}{x^2+y^2+1}.$$

Table 1. The convergence of the method in Example 1

τ	k	$N^{(k)}$	error_1	error_2
0.1	59	36	1.8188e−004	1.1939e−004
0.2	29	36	1.524e−004	1.0195e−004
0.3	18	36	1.8967e−004	1.2542e−004
0.4	13	36	1.7303e−004	1.1205e−004
0.5	10	36	1.4893e−004	1.0241e−004
0.6	8	36	1.1549e−004	1.0472e−004
0.7	9	36	1.055e−004	1.1133e−004
0.8	15	36	1.3817e−004	1.10465e−004
0.9	33	36	1.5119e−004	1.0986e−004

For these data we do not know the exact solution of the problem. Using the proposed method we obtain the results of convergence for $k = 10,\ \tau = 0.5$ given in the Table 2. The graph of the approximate solution is given in Figs. 2.

From the above numerical experiments we see the fast convergence of the domain decomposition method and the efficiency of the truncation method for solving infinite systems of linear equations.

Table 2. The convergence of the method in Example 2

N	$h_1 = h_2 = 1/N$	ε	$N^{(10)}$
16	0.0625	0.1	22
16	0.0625	0.01	35
32	0.03125	0.1	27
32	0.03125	0.01	40

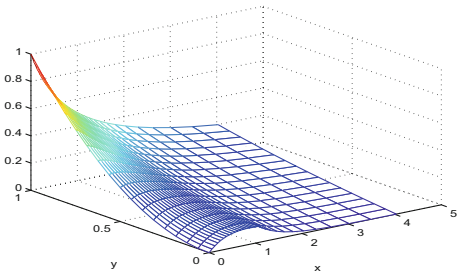


Fig. 2. The graph of the approximate solution with $h_1 = h_2 = 1/16,\ \varepsilon = 0.01,\ \tau = 0.5,\ N^{(10)} = 35$ in Example 2.

5 Conclusion

In this paper, we have proposed and realized a method for solving a strongly mixed elliptic boundary value problem in an unbounded domain. Using the domain decomposition method, we reduce the original problem to two problems, one in a bounded and another in an unbounded subdomain. Next, we reduce infinite systems of three-point vector equations to infinite systems of three-point scalar equations and establish a criterion for determining when truncate the latter infinite systems for assuring to obtain approximate solution with a given accuracy. We used non-equidistant grids with monotonically increasing grid sizes in unbounded subdomain, which significantly reduce the numbers, where the infinite systems are truncated in comparision with uniform grids. Some numerical examples are shown to illustrate the efficiency of the proposed method.

The development of the method for solving other mixed boundary value problems including three-dimensional ones, for elliptic equations in unbounded domains is the direction of our research in the future.

Acknowledgments. This work is supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under the grant number 102.01-2014.20.

References

1. Dang, Q.A., Tran, D.H.: Method of infinite system of equations for problems in unbounded domains. J. Appl. Math. **2012**, 17. doi:10.1155/2012/584704 (2012). Article ID 584704
2. Dang, Q.A., Tran, D.H.: Method of infinite systems of equations for solving an elliptic problem in a semistrip. Appl. Numer. Math. **87**, 114–124 (2015)
3. Dang, Q.A., Truong, H.H.: Simple iterative method for solving problems for plates with partial internal supports. J. Eng. Math. **86**, 139–155 (2014)
4. Dang, Q.A., Truong, H.H.: Numerical solution of the problems for plates on partial internal supports of complicated configurations. J. Phys. Conf. Ser. **490**, 012060 (2014)
5. Dang, Q.A., Vu, V.Q.: A domain decomposition method for strongly mixed boundary value problems for the Poisson equation. In: Bock, H.G., et al. (eds.) Modeling, Simulation and Optimization of Complex Processes, pp. 65–76. Springer, Heidelberg (2012)
6. Duffy, D.G.: Mixed Boundary Value Problems. Taylor & Francis (2008)
7. Fabrikant, V.I.: Mixed Boundary Value Problems of Potential Theory and Their Applications in Engineering. Kluwer Academic Publishers, Dordrecht (1991)
8. Georgiou, G.C., Olson, L.G., Smyrlis, Y.S.: A singular function boundary integral method for the Laplace equation. Comm. Numer. Methods Eng. **12**, 127–134 (1996)
9. Elliotis, M., Georgiou, G., Xenophontos, C.: Solving Laplacian problems with boundary singularities: a comparison of a singular function boundary integral method with the p/hp version of the finite element method. Appl. Math. Comput. **169**, 485–499 (2005)
10. Kantorovich, L.V., Krylov, V.I.: Approximate Methods of Higher Analysis. Phys.-Mat. Publ, Moscow (1962)

11. Li, Z.-C., Chan, Y.L., Georgiou, G., Xenophontos, C.: Special boundary approximation methods for Laplace's equation with singularities. Comput. Math. Appl. **51**, 115–142 (2006)
12. Li, Z.C., Lu, T.T.: Singularities and treatments of elliptic boundary value problems. Math. Comput. Model. **31**, 97–145 (2000)
13. Polozhii, G.N.: The Method of Summary Representations for Numerical Solution of Problems of Mathematical Physics. Pergamon Press, Oxford (1965)
14. Samarskii, A.: The Theory of Difference Schemes. Marcel Dekker, New York (2001)
15. Samarskij, A., Nikolaev, E.: Numerical Methods for Grid Equations, vol. 2. Birkhauser, Basel (1989)
16. Snedon, I.: Mixed Boundary Value Problems in Potential Theory, Amesterdam. North Holl Pub., Com (1966)
17. Vu, V.Q.: The results of using the method of complete reduction for solving elliptic problems with mixed boundary conditions. In: Proceedings of The National Workshop "Development of IT Tools for Teaching, Researching and Applying Mathematics", Hanoi, pp. 247–256 (2005)

Numerical Solution of a Fully Fourth Order Nonlinear Problem

Dang Quang A[1(✉)] and Ngo Thi Kim Quy[2]

[1] Center for Informatics and Computing, Vietnam Academy of Science and Technology (VAST), 18 Hoang Quoc Viet, Cau Giay, Hanoi, Vietnam
dangquanga@cic.vast.vn
[2] Thai Nguyen University of Economic and Business Administration, Thai Nguyen, Vietnam
kimquykttn@gmail.com

Abstract. In this paper we consider a fully fourth order nonlinear boundary value problem which models a statistically bending elastic beam. Differently from the approaches of the other authors, we reduce the problem to an operator equation for the right-hand side function. The existence and uniquenness of solution is established. Besides, we propose an iterative method for solving the problem. Some numerical examples demonstrate the efficiency of the method.

Keywords: Existence and uniqueness of solution · Iterative method · Fully fourth order nonlinear boundary value problem

1 Introduction

In this paper we consider the boundary value problem

$$u^{(4)}(x) = f(x, u(x), u'(x), u''(x), u'''(x)), \quad 0 < x < 1, \tag{1}$$

$$u(0) = u(1) = u''(0) = u''(1) = 0. \tag{2}$$

where $f : [0, 1] \times \mathbb{R}^4$ is continuous.

This problem models the bending equilibrium of a beam on an elastic foundation, whose two ends are simply supported. The special case of Eq. (1), where f does not contain derivative terms u' and u''', i.e.

$$u^{(4)}(x) = f(x, u(x), u''(x)), \quad 0 < x < 1, \tag{3}$$

has studied by several authors. For example, in 1986, Aftabizadeh [1] showed the existence of a solution to the problem (2), (3) under the restriction that f is a bounded function. The result is based on the Leray-schauder fixes point theorem. In 1997 Ma et al. [9] and in 2004 Bai et al. [4] developed the monotone method in the presence of lower and upper solutions for the problem under some monotone conditions of f. The idea of Bai et al. is used in a recent work of Li [7].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_45

Except for the existence Li successfully investigated the uniqueness of the solution to the problem. It should be emphasized that in the monotone method the assumption of the presence of lower and upper solutions is always needed and the finding of them is not easy.

Differently from the approaches to the problem (2), (3) of the authors mentioned above and the approaches to other nonlinear fourth order differential equations with various boundary conditions including nonlocal equations and nonlocal boundary conditions, e.g., in [2,3,10,11], where the problem is led to integral operators for the unknown function $u(x)$, in [6] we reduce the original problem to an operator equation for the right-hand side function. This idea was used by ourselves first in a previous paper [5] when studying the Neumann problem for a biharmonic type equation.

For fully fourth order nonlinear boundary value problem (1)–(2), in 2013, Li and Liang [8] established the existence of solution for the problem under the restriction of the linear growth of the function $f(x, u, y, v, z)$ in each variable on the infinity. In the present paper by a completely different approach we free this restriction. Due to the reduction of the problem to an operator equation for the right hand side function, which will be proved to be contractive, we establish the existence and uniqueness of a solution and the convergence of an iterative method for finding the solution. Some examples demonstrate the applicability of our approach and the efficiency of the proposed iterative method.

2 The Existence and Uniqueness of a Solution

For investigating the problem (1)–(2) we set

$$\varphi = f(x, u, u', u'', u'''), \quad v = u''. \tag{4}$$

Then the problem is reduced to the two second order problems

$$\begin{cases} v'' = \varphi, \quad 0 < x < 1, \\ v(0) = v(1) = 0, \end{cases} \tag{5}$$

$$\begin{cases} u'' = v, \quad 0 < x < 1, \\ u(0) = u(1) = 0, \end{cases} \tag{6}$$

Clearly, the solutions v and u of the above problems depend on φ, that is $v = v_\varphi(x), u = u_\varphi(x)$. Therefore, for φ we have the equation

$$\varphi = A\varphi, \tag{7}$$

where A is a nonlinear operator defined by

$$(A\varphi)(x) = f(x, u_\varphi(x), y_\varphi(x), v_\varphi(x), z_\varphi(x)), \tag{8}$$

with $y_\varphi(x) = u'_\varphi(x), z_\varphi(x) = v'_\varphi(x)$.
We shall prove that under some conditions A is contractive operator.
For each number $M > 0$ denote

$$\mathcal{D}_M = \left\{ (x, u, y, v, z) \mid 0 \le x \le 1, |u| \le \frac{M}{64}, |y| \le \frac{M}{16}, |v| \le \frac{M}{8}, |z| \le \frac{M}{2} \right\}, \tag{9}$$

and by $B[O,M]$ we denote a closed ball centered at O with the radius M in the space of continuous functions $C[0,1]$ with the norm

$$\|\varphi\| = \max_{0\le x\le 1} |\varphi(x)|.$$

Lemma 1. *Assume that there exist numbers $M, c_0, c_1, c_2, c_3 \ge 0$ such that*

$$|f(x,u,y,v,z)| \le M, \tag{10}$$

$$|f(x,u_2,y_2,v_2,z_2,) - f(x,u_1,y_1,v_1,z_1)| \le \\ c_0|u_2-u_1| + c_1|y_2-y_1| + c_2|v_2-v_1| + c_3|z_2-z_1| \tag{11}$$

for any $(x,u,y,v,z),(x,u_i,y_i,v_i,z_i) \in \mathcal{D}_M$ $(i=1,2)$.

Then, the operator A defined by (8), where v_φ, u_φ are the solutions of the problems (5), (6), maps the closed ball $B[O,M]$ into itself. Moreover, if

$$q := \frac{c_0}{64} + \frac{c_1}{16} + \frac{c_2}{8} + \frac{c_3}{2} < 1 \tag{12}$$

then A is contractive operator in $B[O,M]$.

Proof. Let φ be a function in $B[O,M]$. The solution of the problem (5) can be represented in the form

$$v = -\int_0^1 G(x,t)\varphi(t)dt,$$

where $G(x,t)$ is the Green function

$$G(x,t) = \begin{cases} x(1-t), & 0 \le x \le t \le 1, \\ t(1-x), & 0 \le t \le x \le 1. \end{cases}$$

Since

$$\int_0^1 |G(x,t)|dt \le \frac{1}{8}, \quad x \in [0,1] \tag{13}$$

$$\int_0^1 |G'_x(x,t)|dt \le \frac{1}{2}, \quad x \in [0,1] \tag{14}$$

we have

$$\|v\| \le \frac{1}{8}\|\varphi\| \le \frac{1}{8}M. \tag{15}$$

For the solution of the problem (6) we have the estimate

$$\|u\| \le \frac{1}{8}\|v\| \le \frac{1}{64}M. \tag{16}$$

Hence

$$\begin{aligned} \|z\| = \|v'\| \le \frac{1}{2}\|\varphi\| \le \frac{1}{2}M, \\ \|y\| = \|u'\| \le \frac{1}{2}\|v\| \le \frac{1}{16}M. \end{aligned} \tag{17}$$

Therefore, taking into account (8) and the condition (10), we have $A\varphi \in B[0, M]$, i.e. the operator maps $B[0, M]$ into itself.
Now, let $\varphi_1, \varphi_2 \in B[O, M]$ and $v_1, v_2; u_1, u_2$ be the solutions of the problems (5), (6), i.e., for $i = 1, 2$

$$\begin{cases} v_i'' = \varphi_i, \quad 0 < x < 1, \\ v_i(0) = v_i(1) = 0, \end{cases} \tag{18}$$

$$\begin{cases} u_i'' = v_i, \quad 0 < x < 1, \\ u_i(0) = u_i(1) = 0. \end{cases} \tag{19}$$

Using the representation of the solutions v_i and u_i via the Green function and the estimates (13) and (14) we obtain

$$\begin{aligned} &\|v_2 - v_1\| \le \frac{1}{8}\|\varphi_2 - \varphi_1\|, \|u_2 - u_1\| \le \frac{1}{64}\|\varphi_2 - \varphi_1\|, \\ &\|z_2 - z_1\| \le \frac{1}{2}\|\varphi_2 - \varphi_1\|, \|y_2 - y_1\| \le \frac{1}{2}\|v_2 - v_1\| \le \frac{1}{16}\|\varphi_2 - \varphi_1\|. \end{aligned} \tag{20}$$

Now from (8) and (11) it follows

$$\begin{aligned} |A\varphi_2 - A\varphi_1| &= |f(x, u_2, y_2, v_2, z_2) - f(x, u_1, y_1, v_1, z_1)| \\ &\le c_0|u_2 - u_1| + c_1|y_2 - y_1| + c_2|v_2 - v_1| + c_3|z_2 - z_1|. \end{aligned}$$

Using the estimate (20) we obtain

$$\|A\varphi_2 - A\varphi_1\| \le \left(\frac{c_0}{64} + \frac{c_1}{16} + \frac{c_2}{8} + \frac{c_3}{2}\right)\|\varphi_2 - \varphi_1\|$$

Therefore, A is an contractive operator in $B[O, M]$ provided the condition (12) is satisfied. □

Theorem 1. *Under the assumptions of Lemma 1, the problem (1)–(2) has a unique solution u and there hold the estimates*

$$\|u\| \le \frac{M}{64}, \quad \|u'\| \le \frac{M}{16}, \quad \|u''\| \le \frac{M}{8}, \quad \|u'''\| \le \frac{M}{2}. \tag{21}$$

Proof. It is easy to see that the solution of the problem (1)–(2) is the function $u(x)$ obtained from the problems (5), (6), where φ is the unique fixed point of A. The estimates (21) indeed are the estimates (15)–(17). □

3 Iterative Method and Numerical Examples

Consider the following iterative process:

1. Given
$$\varphi_0(x) = f(x, 0, 0, 0, 0). \tag{22}$$
2. Knowing φ_k $(k = 0, 1, ...)$ solve consecutively two problems
$$\begin{cases} v_k'' = \varphi_k(x), \quad 0 < x < 1, \\ v_k(0) = v_k(1) = 0, \end{cases} \tag{23}$$
$$\begin{cases} u_k'' = v_k(x), \quad 0 < x < 1, \\ u_k(0) = u_k(1) = 0. \end{cases} \tag{24}$$
3. Update
$$\varphi_{k+1} = f(x, u_k, u_k', v_k, v_k'). \tag{25}$$

Set $p_k = \dfrac{q^k}{1-q}\|\varphi_1 - \varphi_0\|$. We obtain the following result.

Theorem 2. *Under the assumptions of Lemma 1 the above iterative method converges with the rate of geometric progression and there hold the estimates*
$$\begin{aligned} &\|u_k - u\| \le \frac{p_k}{64}, \quad \|u_k' - u'\| \le \frac{p_k}{16}, \\ &\|u_k'' - u''\| \le \frac{p_k}{8}, \quad \|u_k''' - u'''\| \le \frac{p_k}{2}, \end{aligned} \tag{26}$$
where u is the exact solution of the problem (1)–(2).

Proof. Notice that the above iterative method is the successive iteration method for finding the fixed point of the operator A with the initial approximation (22) belonging to $B[O, M]$. Therefore, it converges with the rate of geometric progression and there is the estimate
$$\|\varphi_k - \varphi\| \le \frac{q^k}{1-q}\|\varphi_1 - \varphi_0\|. \tag{27}$$
Combining it with the estimate of the type (20) we obtain (26), and the theorem is proved.

For numerical realization of the iterative method we use the difference schemes of second order accuracy on uniform grids for the Dirichlet problems (23), (24). The iterations are performed until $e_k = \|u_k - u_{k-1}\| \le 10^{-16}$. In the tables of results of computation n is the number of grid points, $error = \|u_k - u_d\|$, where u_d is the exact solution.

We shall test the convergence of the method for the case of known exact solutions and also for the case of unknown exact solutions.

First, we consider an example for the case of known exact solution.

Example 1. Consider the boundary value problem
$$\begin{cases} u^{(4)}(x) = -\dfrac{u'''(x)}{3} + \cos\left(-\dfrac{\sin \pi x}{\pi^2} - u''(x)\right) - u'(x) - u^2(x) + \sin \pi x \\ \qquad + \dfrac{\cos \pi x}{\pi^3} + \dfrac{\sin^2 \pi x}{\pi^8} - \dfrac{\cos \pi x}{3\pi} - 1, \quad 0 < x < 1, \\ u(0) = u(1) = u''(0) = u''(1) = 0. \end{cases}$$

The exact solution in $[0,1]$ is

$$u(x) = \frac{\sin(\pi x)}{\pi^4}.$$

In this example

$$f(x,u,y,v,z) = -\frac{z}{3} + \cos\left(-\frac{\sin \pi x}{\pi^2} - v\right) - y - u^2 + \sin \pi x + \frac{\cos \pi x}{\pi^3} + \frac{\sin^2 \pi x}{\pi^8} - \frac{\cos \pi x}{3\pi} - 1.$$

Choose M such that $|f(x,u,y,v,z)| \le M$. The number M may be defined from the inequality

$$|f(x,u,y,v,z)| \le \frac{M}{6} + 1 + \frac{M}{16} + \left(\frac{M}{64}\right)^2 + 1 + \frac{1}{\pi^3} + \frac{1}{\pi^8} + \frac{1}{3\pi} + 1 \le M.$$

We can take $M = 6$. Then in the domain $\mathcal{D}_6$, since

$$f'_u = -2u, \quad f'_y = -1, \quad f'_v = \sin\left(-\frac{\sin \pi x}{\pi^2} - v\right), \quad f'_z = -\frac{1}{3},$$

$$|f'_u| \le 2\left(\frac{6}{64}\right) = \frac{3}{16}, \quad |f'_y| \le 1, \quad |f'_v| \le 1, \quad |f'_z| \le \frac{1}{3}$$

we can take $c_0 = \frac{3}{16}, c_1 = c_2 = 1, c_3 = \frac{1}{3}$. Therefore, $q = \frac{c_0}{64} + \frac{c_1}{16} + \frac{c_2}{8} + \frac{c_3}{2} \approx 0.3571 < 1$. All the conditions of Theorem 1 are satisfied. Hence, the problem has a unique solution, and the iterative method converges.

The convergence of the iterative method for Example 1 is given in Table 1 and Fig. 1.

Table 1. The convergence in Example 1.

n	k	$Error$	e_k
30	12	$1.8797e-5$	$3.4694e-18$
50	12	$6.7627e-6$	$3.4694e-18$
100	12	$1.6902e-6$	$8.6736e-19$
1000	12	$1.6901e-8$	$5.2042e-18$

From the Table 1 we observe that the convergence of the iterative method does not depend on the grid size.

In the next example the exact solution of problem (1)–(2) is unknown.

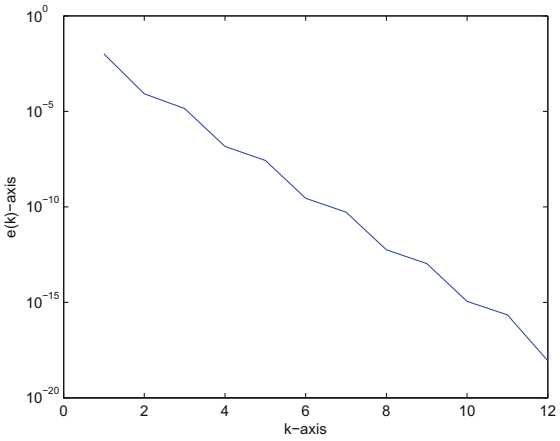


Fig. 1. The graph of e_k in Example 1 for $n = 100$.

Example 2. Consider the boundary value problem

$$\begin{cases} u^{(4)}(x) = -\dfrac{u'''(x)}{12} - u^2(x)\,(u''(x))^2 - (u''(x))^3 + \dfrac{u'(x)}{2} + \sin \pi x + 2, \quad 0 < x < 1, \\ u(0) = u(1) = u''(0) = u''(1) = 0. \end{cases}$$

In this example

$$f(x, u, y, v, z) = -\frac{z}{12} - u^2 v^2 - v^3 + \frac{y}{2} + \sin \pi x + 2.$$

Analogously as in Example 1 we can choose $M = 4$, and therefore, the Lipschitz coefficients in Lemma 1 are $c_0 = 0.03, c_1 = \dfrac{1}{2}, c_2 = 0.75, c_3 = 0.083$. Then

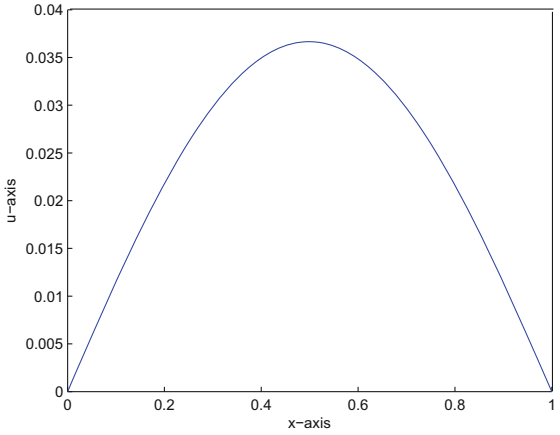


Fig. 2. The graph of the approximate solution in Example 2.

$q = \frac{c_0}{64} + \frac{c_1}{16} + \frac{c_2}{8} + \frac{c_3}{2} \approx 0.167 < 1$. Hence, the problem has a unique solution, and the iterative method converges.

Remark that the function $f(x, u, y, v, z)$ does not satisfy the conditions in [8], so, the theorems of that paper cannot ensure the existence of a solution.

The numerical experiment for $n = 100$ shows that with the above stopping criterion after $k = 9$ iterations the iterative process stops and $e_9 = 4.5103e - 17$.

The graph of the approximate solution in Example 2 is depicted in Fig. 2.

4 Conclusion

In this paper we have proposed a method for investigating the solvability and iterative solution of a nonlinear fully fourth order boundary value problem. In the method by the reduction of the problem to an operator equation for the right hand side we have established the existence and uniqueness of a solution and the convergence of an iterative process.

The proposed method can be used for some other problems for ordinary and partial differential equations.

Acknowledgments. This work is supported by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under the grant number 102.01-2014.20.

References

1. Aftabizadeh, A.R.: Existence and uniqueness theorems for fourth-order boundary value problems. J. Math. Anal. Appl. **116**, 415–426 (1986)
2. Alves, E., Ma, T.F., Pelicer, M.L.: Monotone positive solutions for a fourth order equation with nonlinear boundary conditions. Nonlinear Anal. **71**, 3834–3841 (2009)
3. Amster, P., Cárdenas Alzate, P.P.: A shooting method for a nonlinear beam equation. Nonlinear Anal. **68**, 2072–2078 (2008)
4. Bai, Z., Ge, W., Wang, Y.: The method of lower and upper solutions for some fourth-order equations. J. Inequalities Pure Appl. Math. **5**(1), 1–17 (2004). Article no. 13
5. Dang, Q.A.: Iterative method for solving the Neumann boundary value problem for biharmonic type equation. J. Comput. Appl. Math. **196**, 634–643 (2006)
6. Dang, Q.A., Dang, Q.L., Ngo, T.K.Q.: A novel efficient method for fourth order nonlinear boundary value problems, submitted to Numerical Algorithms
7. Li, Y.: A monotone iterative technique for solving the bending elastic beam equations. Appl. Math. Comput. **217**, 2200–2208 (2010)
8. Li, Y., Liang, Q.: Existence results for a fourth-order boundary value problem. J. Funct. Spaces Appl. **2013**, 5 (2013). Article ID 641617
9. Ma, R., Zhang, J., Fu, S.: The method of lower and upper solutions for fourth-order two-point boundary value problems. J. Math. Anal. Appl. **215**, 415–422 (1997)
10. Ma, T.F., da Silva, J.: Iterative solutions for a beam equation with nonlinear boundary conditions of third order. Appl. Math. Comput. **159**, 11–18 (2004)
11. Ma, T.F.: Existence results and numerical solutions for a beam equation with nonlinear boundary conditions. Appl. Numeri. Math. **47**, 189–196 (2003)

On the Performance of Decode-and-Forward Half-Duplex Relaying with Time Switching Based Energy Harvesting in the Condition of Hardware Impairment

Tan N. Nguyen[1,2], Phuong T. Tran[1(✉)], Huong-Giang Hoang[1], Hoang-Sy Nguyen[2], and Miroslav Voznak[2]

[1] Faculty of Electrical and Electronics Engineering, Ton Duc Thang University, 19 Nguyen Huu Tho Street, District 7, Ho Chi Minh City, Vietnam
{nguyennhattan,tranthanhphuong,hoangthihuonggiang}@tdt.edu.vn
[2] VSB-Technical University of Ostrava, 17. listopadu 15/2172, 708 33 Ostrava, Poruba, Czech Republic
{sy.nguyen.hoang.st,miroslav.voznak}@vsb.cz
http://dee.tdt.edu.vn

Abstract. In this paper, the outage performance and system throughput of a decode-and-forward half-duplex relay network, in which the relay node is equipped with time switching based energy harvesting capability, are derived rigorously. The analytical results provide theoretical insights into the effect of various system parameters, such as time switching factor, source transmission rate, transmitting-power-to-noise ratio to system performance. Our analysis is confirmed by Monte-Carlo simulation, and can serve as a guideline to design practical energy-harvested relaying systems.

Keywords: Hardware impairment · Energy harvesting · Time switching · Decode-and-forward · Half-duplex relaying

1 Introduction

Relaying is an effective way to combat the performance degradation caused by fading, shadowing, and path loss. However, conventional relay nodes are powered by batteries, which have a limited operational lifetime, and have to be replaced or recharged periodically. This drawback limits the benefit of relaying in wireless networks in general. Recently, simultaneous wireless information and power transfer (SWIPT) has been presented in several studies, starting with the pioneering works by Varshney [1] and Grover [2]. SWIPT is a kind of communications where relay nodes can be powered by energy harvested from the source RF signals.

Since then, RF energy harvesting for relay networks has attracted significant attention from researchers. Time-switching and power-splitting protocol

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_46

for energy harvesting have been introduced in [3] as two practical architectures for SWIPT in relay networks. However, this work only shown a theoretical upper bound on system performance, and was lack of practical consideration. Later, Nasir et al. [4,5] shed light into the effect of different system parameters on the throughput performance of amplify-and-forward (AF) and decode-and-forward (DF) relaying systems for both time-switching (TSR) and power-splitting (PSR) protocols.

Common to all these works studying energy harvesting in relay networks is the assumption of perfect transceiver hardware at all nodes. However, in practice, the transceiver hardware is imperfect due to phase noise, I/Q imbalance and amplifier nonlinearities [6]. In [7], the authors have shown that hardware impairments limit the performance of dual-hop relaying systems, in terms of the capacity, throughput and symbol error rate (SER). The joint impact of hardware impairment and co-channel interference on performance of relay networks was also mentioned in [8].

In this paper, we take account of the impact of hardware impairments at both source and relay nodes in the analysis of outage probability and throughput performance of wireless information and power transfer with half-duplex relaying. Regarding to the relaying and energy harvesting protocols, we focus on decode-and-forward relaying and time-switching energy harvesting protocol. The effect of hardware impairments on system performance systems is evaluated both by mathematical analysis and simulation. The impact of other parameters such as power supply and time switching factor on system performance under hardware impairments is also investigated. Furthermore, the optimal time-switching factor to maximize system throughput is found by numerical search algorithm.

The rest of this paper is organized as follows. The system model is described in Sect. 2. In Sect. 3, closed-form expressions for the outage probability and average throughput are derived. Numerical results to confirm the mathematical analysis are presented in Sect. 4. Finally, this paper is concluded in Sect. 5.

2 System Model

We consider a half-duplex relaying system as illustrated in Fig. 1, where the source S sends information to the destination D with the help of a decode-and-forward relay R. Assume that the direct connection between source and destination is weak, so the relay R is deployed to enhance this connection without having its own data to transmit. The relay is assumed to have no other energy supply but only the energy harvested from the source.

For separating between information transferring and energy harvesting processes at relay node, we adopt the time-switching relaying protocol (TSR) [4]. Here, the total symbol duration T is divided into two intervals with the lengths of α and $1-\alpha$, respectively. The first interval corresponds to the energy harvesting phase and the second one corresponds to the information transmission phase. The latter interval is then divided into two subintervals with equal length. In the first subinterval, relay node receives the signal from the source and decodes

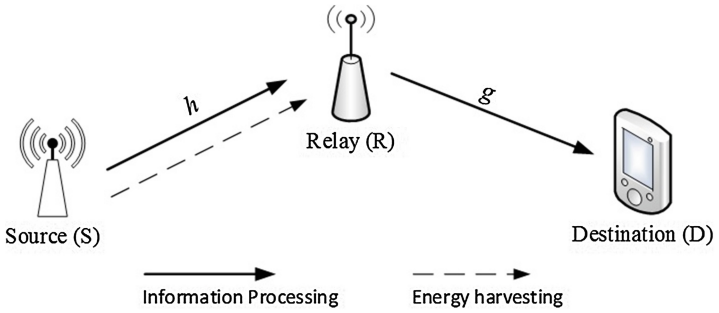


Fig. 1. System model.

it. The decoded message is then forwarded to the destination during the second subinterval. Figure 2 explains the TSR protocol visually.

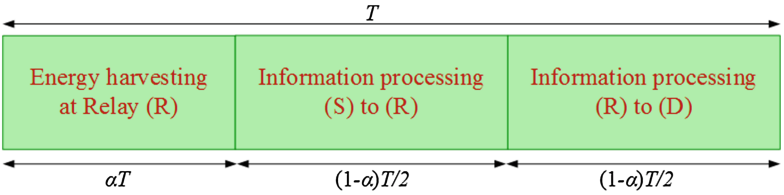


Fig. 2. Time-switching relaying protocol.

Regarding to the channel model, we consider the case that the channel state information is obtained without error. Let h and g denote the channels from the source to the relay and from the relay to the destination, respectively. We assume that all channels experience Rayleigh fading and keep constant during each transmission block (flat fading). As a result, $|h|^2$ is an exponential random variable with mean λ_h, and $|g|^2$ is exponentially distributed with mean λ_g.

2.1 Energy Harvesting Phase

As mentioned above, the relay carries out the energy capturing process during the energy harvesting phase. The received signal at the relay node in this phase can be expressed as

$$y_{re} = hx_e + n_r \tag{1}$$

where h is the channel coefficient of the source-relay path, x_e is the energy-transmitted signal with $E|x_e|^2 = P_s$, where $E[.]$ denotes the expectation operation. n_r is the zero-mean additive white Gaussian noise (AWGN) with variance N_0. Hence, the energy harvested during this phase can be computed as

$$E_h = \mu P_s |h|^2 \alpha T \tag{2}$$

where μ is a constant and denotes the energy conversion efficiency. We ignore the hardware noise term because we only care about the energy of signals. The impact of noise, if exists, just add more energy to the receive signal at R.

The relay transmit power can be computed as [9]

$$P_r = \frac{E_h}{(1-\alpha)T/2} = \frac{2\mu P_s |h|^2 \alpha}{(1-\alpha)} = 2kP_s|h|^2 \tag{3}$$

where k is defined as $k = \frac{\mu\alpha}{1-\alpha}$.

2.2 Information Transmission Phase

Now, we consider the information transmission phase. This phase is divided into two equal-length subintervals. In the first interval, due to hardware impairment, the received signal at the relay can be expressed as

$$y_r = h(x_s + \eta_s) + n_r \tag{4}$$

where x_s is the transmitted signal, which satisfies $E|x_s|^2 = P_s$, η_s denotes the distortion error caused by hardware impairment at source node, which is modeled as a zero-mean Gaussian random variable with variance $P_s\sigma_1^2$, and n_r denotes the AWGN noise at relay node. As a result, the signal-to-impairment-noise-ratio (SINR) at relay node can be calculated by the following equation:

$$\gamma_1 = \frac{P_s|h|^2}{P_s|h|^2{\sigma_1}^2 + N_0} \tag{5}$$

where N_0 is the noise power at the relay.

During the second interval, the destination is receiving the forwarded signal from the relay. We assume that the link between source and destination is very weak, so the communication in this interval is relied almost on the relay-destination link. Hence, the received signal at destination node is given by

$$y_D = g(x_r + \eta_r) + n_d \tag{6}$$

where x_r is the signal transmitted by the relay which satisfies the constraint $E|x_r|^2 = P_r$, η_r is the noise caused by impairment at relay node, and n_d is the noise at the destination, which is assumed to have the same power as n_r. Now, the SINR at the destination node can be calculated by

$$\gamma_2 = \frac{P_r|g|^2}{P_r|g|^2{\sigma_2}^2 + N_0} = \frac{2kP_s|h|^2|g|^2}{2kP_s|h|^2|g|^2{\sigma_2}^2 + N_0} \tag{7}$$

where σ_2 denotes the level of impairments at relay node, i.e., $P_r\sigma_2^2$ is the variance of η_r.

3 Performance Analysis

In this section, we proceed to analyze the outage and throughput performance of two-hop DF half-duplex relaying network with RF energy harvesting at the relay. The dependence of average throughput as well as the outage probability of the proposed system on the time-switching factor is derived and the optimal time allocation is found. Here, we consider the instantaneous transmission mode.

3.1 Outage Probability

Let's derive the formula for the outage probability of the proposed model first. Assume that the source transmits at a constant rate R. Let $z = 2^R - 1$ be the lower threshold for SINR at both relay and destination nodes. In addition, all transmitters are assumed to have the same hardware structure, so that the impairments level are the same, i.e., $\sigma_1 = \sigma_2 = \sigma$. For DF relaying protocol, the data communication is divided into two separating hops, which do not depend on each other. Hence, the outage occurs if and only if either the source-relay path or the relay-destination path fails to satisfy the corresponding SINR constraint. That means the outage probability of the system can be written as

$$P_{out} = \Pr\left(\min\left(\gamma_1, \gamma_2\right) < z\right) = 1 - \Pr\{\gamma_1 \geq z, \gamma_2 \geq z\} \tag{8}$$

Then we can claim the following theorem on the outage probability of the system of interest.

Theorem 1. *For the DF half-duplex relaying system with time-switching based energy harvesting that satisfies the condition $z\sigma^2 < 1$, the outage probability of the system can be expressed as*

$$P_{out} = 1 - e^{-\frac{x_0}{\lambda_h} - \frac{1}{2k\lambda_g}} - \frac{1}{\lambda_g} \int_0^{\frac{1}{2k}} e^{-\frac{y}{\lambda_g} - \frac{x_0}{2\lambda_h k y}} dy \tag{9}$$

where

$$x_0 = \frac{N_0 z}{P_s(1 - \sigma^2 z)} \tag{10}$$

Proof. Let's denote $X = |h|^2$ and $Y = |g|^2$, then X and Y are exponential random variables with parameters λ_h and λg, respectively. The Eq. (8) can be rewritten as

$$\begin{aligned} P_{out} &= 1 - \Pr\left(\frac{P_s X}{P_s X \sigma^2 + N_0} \geq z, \frac{2kP_s XY}{2kP_s XY\sigma^2 + N_0} \geq z\right) \\ &= 1 - \Pr\left(P_s X(1 - z\sigma^2) \geq N_0 z, 2kP_s XY(1 - z\sigma^2) \geq N_0 z\right) \end{aligned} \tag{11}$$

In the following analysis, we assume that the hardware impairment is limited so that the condition $z\sigma^2 \leq 1$ is satisfied. Now, (11) can be rewritten as

$$P_{out} = 1 - \Pr\left(X \geq \frac{N_0 z}{P_s(1-z\sigma^2)}, XY \geq \frac{N_0 z}{2kP_s(1-z\sigma^2)}\right) \tag{12}$$

By letting $x_0 = \frac{N_0 z}{P_s(1-\sigma^2 z)}$ and using the independence of random variables X and Y (which implies $f_{XY}(x,y) = f_X(x).f_Y(y)$, where $f_{XY}(\cdot,\cdot), f_X(\cdot)$, and $f_Y(\cdot)$ are the joint pdf of X and Y, the pdf of X, and the pdf of Y, respectively), the above equation can be reduced to

$$\begin{aligned} P_{out} &= 1 - \Pr(X \geq x_0, XY \geq \frac{x_0}{2k}) = 1 - \iint\limits_{S=\left\{x \geq x_0, xy \geq \frac{x_0}{2k}\right\}} f_X(x) f_Y(y) dx dy \\ &= 1 - \int\limits_0^{y_0} f_Y(y) dy \int\limits_{x_0/(2ky)}^{\infty} f_X(x) dx - \int\limits_{y_0}^{\infty} f_Y(y) dy \int\limits_{x_0}^{\infty} f_X(x) dx \\ &= 1 - e^{-\frac{x_0}{\lambda_h} - \frac{1}{2k\lambda_g}} - \frac{1}{\lambda_g} \int\limits_0^{\frac{1}{2k}} e^{-\frac{y}{\lambda_g} - \frac{x_0}{2\lambda_h k y}} dy \end{aligned} \qquad \square$$

Remark 1. In this analysis, we assume that the variance of hardware impairment error is upper-bounded, so that the quantity $1 - \sigma^2 . z$ is positive, otherwise, the outage probability would go to 1. This is reasonable in practice, because we cannot increase the rate arbitrarily. For each required transmission rate, there is a corresponding constraint on hardware impairment error.

3.2 System Throughput

Here, the average throughput can be computed in terms of the outage probability as [4]

$$R_{DL}(\alpha) = (1 - P_{out}) R \frac{(1-\alpha)}{2} \tag{13}$$

Thus we have the second theorem stated as follows.

Theorem 2. *For the DF half-duplex relaying system with time-switching based energy harvesting, the outage probability of the system can be expressed as*

$$R_{DL}(\alpha) = \frac{R.(1-\alpha)}{2} . \left(e^{-\frac{x_0}{\lambda_h} - \frac{1}{2k\lambda_g}} + \frac{1}{\lambda_g} \int\limits_0^{\frac{1}{2k}} e^{-\frac{y}{\lambda_g} - \frac{x_0}{2\lambda_h k y}} dy \right) \tag{14}$$

Proof. This result follows by substituting (9) into (13). $\square$

3.3 Optimal Time-Switching Factor

The optimal time-switching factor α^* could be obtained by solving the following optimization problem

$$\alpha^* = \arg\max_{\alpha} R_{DL}(\alpha) \tag{15}$$

This can be done by solving the equation $\frac{dR_{DL}(\alpha)}{d\alpha} = 0$. Although this optimization problem does not admit a closed-form solution, we can solve it numerically by using Golden section search algorithm [10], for example.

4 Numerical Results

In this section, we conduct Monte Carlo simulation to verify the analysis developed in the previous section. For simplicity, in our simulation model, we assume that the source-relay and relay-destination distances are both normalized to unit value. Other simulation parameters are listed in Table 1.

Table 1. Simulation parameters

Symbol	Name	Values
R	Source rate	3 bps/Hz
z	SINR threshold	7
μ	Energy harvesting efficiency	0.8
λ_h	Mean of $\|h\|^2$	0.5
λ_g	Mean of $\|g\|^2$	0.5
P_s/N_0	Signal to noise ratio	0–30 dB

Figures 3 and 4 shows the achievable throughput and outage probability of the system versus P_s/N_0 ratio with the variance of hardware impairment error set to be 0.1. The time-switching factor α is set to 0.3. It's can be observed that the simulation curve and the analytical curve almost overlap together.

This confirms our mathematical analysis in the previous section. We can notice that in the low SINR regime, the hardware impairment error has severe effect on system performance. However, when the transmit power goes large, the hardware impairment error has decreasing impact on both achievable throughput as well as outage probability.

The impact of time switching factor on the instantaneous capacity is shown in Fig. 5. In this experiment, P_s/N_0 is set to 30 dB. Again, the analytical solutions are in exact agreement with the simulation results. There exists a unique time switching factor at which the system throughput is maximized. In practice, this optimal factor can be found iteratively using numerical methods. The outage probability also depends on the time switching factor, which is illustrated in Fig. 6.

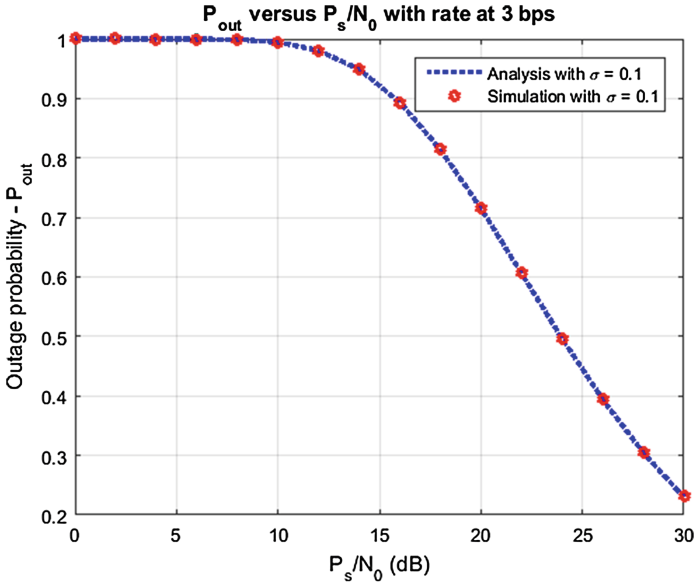


Fig. 3. Outage probability versus source transmit power

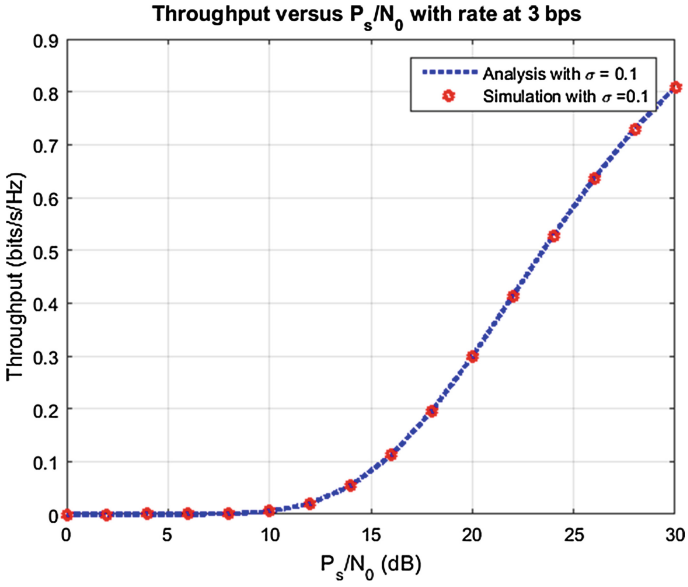


Fig. 4. Throughput versus source transmit power

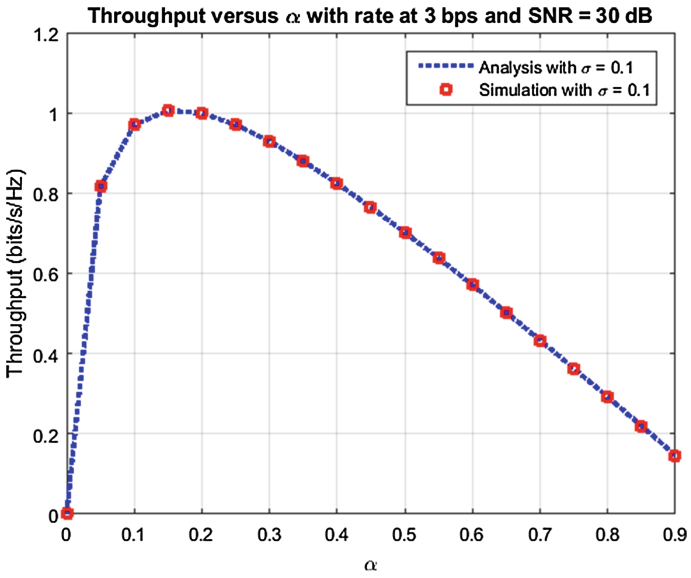


Fig. 5. Outage probability versus time-switching factor

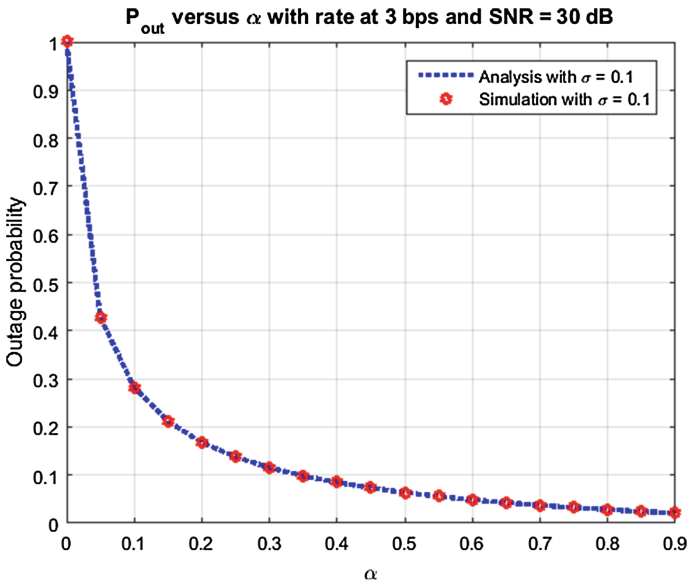


Fig. 6. Throughput versus time-switching factor

5 Conclusion

In this paper, we have studied the performance of DF half-duplex relay networks using RF energy harvesting with time-switching protocol. Distinguished from other studies on RF energy harvesting in relay networks, this paper concentrates on the impact of hardware impairments on the system performance. The hardware impairments at both source and relay nodes, which are modeled as Gaussian random variables, are taken into account in our work. Specifically, we derive the analytical formula for outage probability and system throughput and verify these results by Monte Carlo simulation. From the numerical results, it can be observed that the hardware impairments have severe impact on the system performance at low SINR regime. However, at high SINR regime, the system can still achieve good performance if the impairments are bounded, which is in fact probable in practice.

References

1. Varshney, L.R.: Transporting information and energy simultaneously. In: 2008 IEEE International Symposium on Information Theory, Toronto, ON, Canada, pp. 1612–1616 (2008)
2. Grover, P., Sahai, A.: Shannon meets Tesla: wireless information and power transfer. In: 2008 IEEE International Symposium on Information Theory, Austin, TX, USA, pp. 2363–2367 (2008)
3. Zhang, R., Ho, C.K.: MIMO broadcasting for simultaneous wireless information and power transfer. IEEE Trans. Wireless Commun. **12**(5), 1989–2001 (2013)
4. Nasir, A.A., Zhou, X., Durrani, S., Kennedy, R.A.: Relaying protocols for wireless energy harvesting and information processing. IEEE Trans. Wireless Commun. **12**(7), 3622–3636 (2013)
5. Nasir, A.A., Zhou, X., Durrani, S., Kennedy, R.A.: Throughput and ergodic capacity of wireless energy harvesting based DF relaying network. In: 2014 IEEE International Conference on Communications (ICC), Canberra, ACT, Australia, pp. 4066–4071 (2014)
6. Schenk, T.: RF Imperfections in High-Rate Wireless Systems: Impact and Digital Compensation. Springer, The Netherlands (2008)
7. Bjornson, E., Matthaiou, M., Debbah, M.: A new look at dual-hop relaying: performance limits with hardware impairments. IEEE Trans. Commun. **61**(11), 4512–4525 (2013)
8. Duy, T.T., Duong, T.Q., Da Costa, D.B., Bao, V.N.Q., Elkashlan, M.: Proactive relay selection with joint impact of hardware impairment and co-channel interference. IEEE Trans. Commun. **63**(5), 1594–1606 (2015)
9. Zhou, X., Zhang, R., Ho, C.K.: Wireless information and power transfer: architecture design and rate-energy tradeoff. IEEE Trans. Commun. **61**(11), 4754–4767 (2013)
10. Chong, E.K.P., Zak, S.H.: An Introduction to Optimization, 4th edn. Wiley, New York (2009)

Opportunistic Multiple Relay Selection Schemes in both Full-Duplex and Half-Duplex Operation for Decode-and-Forward Cooperative Networks

Hoang-Sy Nguyen[1,2(✉)], Anh-Hoa Thi Bui[2], Nhat-Tan Nguyen[1,3], and Miroslav Voznak[1]

[1] VSB-Technical University of Ostrava, Ostrava, Czech Republic
{sy.nguyen.hoang.st,miroslav.voznak}@vsb.cz
[2] EIU-Eastern International University, Thu Dau Mot, Binh Duong Province, Vietnam
hoa.bui.k1set@eiu.edu.vn
[3] TDT-Ton Duc Thang University, Ho Chi Minh City, Vietnam
nguyennhattan@tdt.edu.vn

Abstract. An efficient technique used to prolong the lifetime of energy-constrained networks is energy harvesting (EH). In this paper, we investigate and develop the energy allocation methods for the relaying networks, in which opportunistic multiple relay selection schemes with both the full-duplex (FD) and half-duplex (HD) scheme for both EH and non-EH in decode-and-forward (DF) relaying mode. In addition, there are two policies proposed in this paper: (1) Max-Min with Self Interference Relay Selection (MMSI); (2) Max-Min Relay Selection (MMSR) are depicted for both EH and Non-EH relaying modes. Particularly, we derive closed-form expressions of outage probability to analyze the performance of systems. In addition, we propose the impact of self-interference on both policies to provide a practical insight. The results in numerical analysis reveal that the proposed MMSI scheme outperforms the MMSR mode in terms of outage probability.

Keywords: Energy harvesting · Full-duplex · Half-duplex · Decode-and-forward · Power splitting-based relaying · Select relay

1 Introduction

Due to the constantly increasing data rates, the need to utilize wireless energy harvesting emerges, though that proves to be technically difficult. Based on research into enhancing a communications system with delay-limited characteristics, with the purpose to avail the ambient RF (Radio Frequency) signal and harvest energy as well as process information at the same time, a solution emerges in [1–3]. Relay nodes need to be more long-lasting for use in wireless networks in [4,5]. Two given protocols are TSR (Time switching-based relaying) and PSR (Power splitting-based relaying). Furthermore, in [6] we described the cooperative energy harvesting

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_47

communication networks, in [7] the capabilities of MPR (Multipacket Reception) at receivers are portrayed. In addition, SWIPT (simultaneous wireless information and power transfer) network [8] two FD relays and a source - destination pair transmit messages to the target. A promising solution for providing convenient and perpetual energy supplies for low-power wireless sensor networks in particular and for (multiple-input-multiple-output) MIMO networks in general [9–11]. With RF energy from the system called multiuser MIMO, energy harvesting performance of a wireless sensor node is evidenced in [9].

There is a problem with relay selection between the AF and FD communication operations [12]. The best relay selection will maximize instantaneous capacity of FD channel, which requires CSI (channel state information) global channel state, and more importantly several less-than-standard policies of relay selection with the aid of partial CSI knowledge including (a) links between source and relay, relay and destination (b) the source-relay link and interference of loops (c) the loop interference. In [13], OPA (optimal power allocation affected by SPC (individual power constraints) together with SPC (sum power constraints) is presented. Reference [14] establishes the scheme of power assignment in networks of DF (decode and forward) with multi-relay feature over Rayleigh fading channels. In [15] the authors study evidences the best resource allocation for wireless relay networks with multiple pairs beaming from source to destination.

In this paper, we will investigate and derive opportunistic multiple relay selections with both FD and HD transmission modes for DF relaying. The main contributions of this paper are summed up as below:

- This paper derive closed-form expressions of the outage probability for multiple-relay system over independent and identically distributed.
- An important point in this paper is that the outage probability for both MMSI and MMSR policies are illustrated for both energy harvesting (EH) and non-energy harvesting (Non-EH).
- Numerical results reveal the impact of MMSI and MMSR on the performance of systems which includes outage probability.

Notations: $\mathbb{E}\{.\}$ denotes expectation operator, $|.|$ the absolute value operator, $\Pr\{.\}$ is the probability, and $f_X\{.\}$ is probability density functions (PDF), $F_X\{.\}$ is cumulative distribution functions (CDF) of the random variables X.

2 System Model

The following subsections will analyze the harvested energy and information processing at the relay for the PSR protocol. The received signal is defined by P and the whole block time is T. In the half of time, the information transmission uses $T/2$ from source to relays, and another $T/2$ is utilized for data transmission from relays to destination. Meanwhile in the first $T/2$ harvested energy and the remaining received power use βP and $(1-\beta)P$ is utilized for from source to relays information processing with $0 \leq \beta \leq 1$. In Fig. 1 a cooperative relaying network including a source node S, a cluster R_i of N relays ($1 \leq R_i \leq N$) and a destination D.

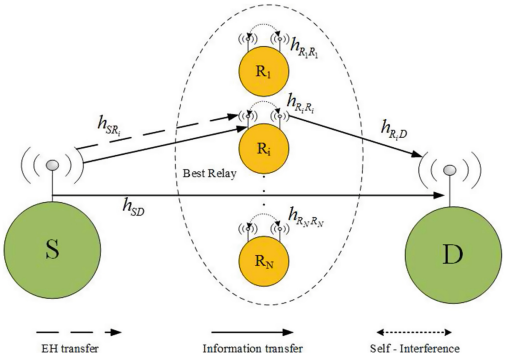


Fig. 1. System model.

There are four wireless links in the system such as source-relay (SR), residual seft-interference (RR), relay-destination (RD), and source-destination (SD) channels shown respectively by $h_{SR_i}, h_{R_iR_i}, h_{R_iD}$ and h_{SD}. Two modeled nature of the channels are frequency-flat and quasi-static. P_S is the normalized transmit power's mark in the source and P_R is the normalized transmit power's mark in the relay, and subject to individual constraints, i.e., $P_S \leqslant 1$ and $P_R \leqslant 1$.

Therefore the received signals in the relay are depicted by

$$y_{R_i} = h_{SR_i} x_S(t) + h_{R_iR_i} x_{R_i}(t) + n_{R_i} \tag{1}$$

where $x_S(t)$ is the information symbol from the source at time slot t, and get $\mathbb{E}\left\{|x_S(t)|^2\right\} = P_S$. $x_R(t)$ is the interference of loopback because of FD relaying and match $\mathbb{E}\left\{|x_R(t)|^2\right\} = P_R$, and n_R means additive white Gaussian noises (AWGN) with $\sigma^2_{n_{R_i}}$ the variances of the AWGNs. The received signals in the destination are presented as

$$y_D = h_{R_iD} x_{R_i}(t) + n_D, \tag{2}$$

where n_D means AWGN at the destination and σ^2_D the variances of the AWGNs.

We represent the system with channel in terms of a parameter, whose definition are $\gamma_{SR_i} = |h_{SR_i}|^2/\sigma^2_{R_i}$, $\gamma_{R_iD} = |h_{R_iD}|^2/\sigma^2_D$ and $\gamma_{R_iR_i} = |h_{R_iR_i}|^2/\sigma^2_{R_i}$ denoting the difference of residual self-interference to-noise ratio at the i-th relay. In this paper, values of σ^2_D, σ^2_S, σ^2_R with $\sigma^2_D = \sigma^2_S = \sigma^2_R = \sigma^2$ are supposed. Due to the cellular down-link transmission, the equivalent signal-to-noise ratio (SNR) of the chosen link can hence be depicted as

$$\gamma_{eq} = \min\{\gamma_{Ri}, \gamma_D\}, \tag{3}$$

where $\gamma_{R_i} = \frac{P_S\gamma_{SR_i}}{P_R\gamma_{R_iR_i}+N_0} \approx \frac{P_S\gamma_{SR_i}}{P_R\gamma_{R_iR_i}}$ and $\gamma_D = P_R\gamma_{R_iD}$.

3 Performance Analysis

3.1 Outage Probability

According a simple order statistic result, it is possible to write the outage probability of the seperated power constraints (SPC) and EH SPC which are defined by

$$\begin{aligned} F_{\gamma_{eq}}^{FD}(\gamma_0) &= \Pr\{\gamma_{eq} < \gamma_0\} = \Pr\{\gamma_{eq,1} < \gamma_0, \gamma_{eq,2} < \gamma_0, ..., \gamma_{eq,N} < \gamma_0\} \\ &= \prod_{i=1}^{N} \Pr\{\gamma_{eq,i} < \gamma_0\} = [\Pr\{\gamma_{eq,i} < \gamma_0\}]^N, \end{aligned} \quad (4)$$

where the threshold γ_0 is related to the target rate R_0 (bps/Hz). When the cooperative networks work in FD mode, we get $\gamma_0 = 2^{R_0} - 1$. In HD mode, we get $\gamma_0 = 2^{2R_0} - 1$.

3.2 Throughput

In this section, we present the throughput of delay limited mode, τ is communicated by the transmitter and the effective communication time from source to destination is $T/2$. The throughput achieved at the destination node is expressed as

$$\tau = (1 - P_{out})\, R_0/2, \quad (5)$$

where the outage probability, P_{out} of EH and Non-EH relay selection policies will be depicted as follow.

4 Relay Selection Policies

In this section, we depicts two relay selection policies, namely MMSI and MMSR in both EH and Non-EH, which implies the application of optimal relay selection policy in some scenarios. Firstly, we present the vital parts in detail to analyze the system performance with EH.

4.1 Energy Harvesting (EH)

Received signal at the relay in energy harvesting networks are depicted by [16]

$$P_R = \frac{E_h}{T/2} = \frac{\eta\beta T/2 \frac{P_S |h_{SR_i}|^2}{\sigma^2}}{T/2\sigma^2} = \rho\gamma_{SR_i}, \quad (6)$$

where $\rho = \eta\beta P_S \sigma^2$. Hence the outage probability for FD mode of EH SPC assisted network is

$$P_{out}^{EH_SPC}(\gamma_0) = 1 - \left(1 - e^{\left(-\frac{(1-\beta)}{\rho\gamma_0\bar{\gamma}_{RR}}\right)}\right) \cdot \sqrt{\frac{4\gamma_0}{\rho\bar{\gamma}_{SR}\bar{\gamma}_{RD}}} K_1\left(\sqrt{\frac{4\gamma_0}{\rho\bar{\gamma}_{SR}\bar{\gamma}_{RD}}}\right), \quad (7)$$

where the advantages of the exponential random variables γ_{RD}, γ_{SR} and γ_{RR} are $\bar{\gamma}_{RD}$, $\bar{\gamma}_{SR}$ and $\bar{\gamma}_{RR}$, respectively.

Proof: See Appendix A.

A. Max-Min with Self-interference Relay Selection (MMSI): In the opportunistic relay-selection model, the end-to-end mutual information of the FD cooperative system comprising multiple DF relays depends on the selection k relay of the strongest bottleneck link, as given by

$$k_{MMSI} = \arg\max_{i} \min\left\{\gamma_{Ri}, \gamma_D\right\} \tag{8}$$

Therefore, the corresponding outage probability for the high SNR regime can be written as from Eq. (7)

$$F_{MMSI}^{EH_SPC}\left(\gamma_0\right) = \left[P_{out}^{EH_SPC}(\gamma_0)\right]^N. \tag{9}$$

B. Max-Min Relay Selection (MMSR): The conventional max-min (MM) relay selection policy takes no account of the loop interference and the optimal relay selection for conventional HD relaying systems [17].

The MM scheme chooses the relay with the best end-to-end link and relay where

$$k_{MMSR} = \arg\max_{i} \min\left\{\gamma_{R_i}, \gamma_D\right\} \tag{10}$$

In case the system has no impact on loop-interference, it becomes HD relay networks. The end-to-end SNR at i-th relay and destination can be expressed as

$$\begin{cases} \gamma_{Ri} = (1-\beta)\, P_S \gamma_{SRi} \\ \gamma_D = P_R \gamma_{RiD} = \rho \gamma_{SRi} \gamma_{RiD} \end{cases} \tag{11}$$

In this case, the best relay selection scheme brings higher performance by

$$F_{MMSR}^{EH_SPC_HD}\left(\gamma_0\right) = \left[P_{out}^{EH_SPC_HD}\left(\gamma_0\right)\right]^N, \tag{12}$$

where $P_{out}^{EH_SPC_HD}\left(\gamma_0\right) = 1 - \left(1 - e^{-\left(\frac{\gamma_0}{(1-\beta)P_S\bar{\gamma}_{SR}}\right)}\right) \cdot \chi K_1\left(\chi\right)$ and N is the number of relay, $\chi = \sqrt{4\gamma_0(\rho\bar{\gamma}_{RD}\bar{\gamma}_{SR})^{-1}}$.

Proof: See in Appendix B.

4.2 Non Energy Harvesting (Non-EH)

I. Non-EH equipped with the relaying system: Optimal power distribution of the source and relay is given by

$$\begin{cases} \frac{P_S \gamma_{SR_i}}{P_R \gamma_{R_i R_i} + 1} = P_R \gamma_{R_i D} \\ P_S = 1 \end{cases} \tag{13}$$

where $0 \leq (P_S, P_R) \leq 1$. We can gain the average optimal transmit power by using the previous formula, thus

$$\begin{cases} P_R = \min\left\{1, \frac{-\bar{\gamma}_{RD} + \sqrt{\bar{\gamma}_{RD}^2 + 4\bar{\gamma}_{RR}\bar{\gamma}_{SR}\bar{\gamma}_{RD}}}{2\bar{\gamma}_{RR}\bar{\gamma}_{RD}}\right\} \\ P_S = 1 \end{cases} \tag{14}$$

It is possible to calculate the outage probability for FD mode of SPC assisted network as

$$P_{out}^{SPC}(\gamma_0) = 1 - \frac{P_S\bar{\gamma}_{SR}}{P_S\bar{\gamma}_{SR} + P_R\bar{\gamma}_{RR}\gamma_0} \frac{1}{\exp\left[\left(\frac{1}{P_R\bar{\gamma}_{RD}} + \frac{1}{P_S\bar{\gamma}_{SR}}\right)\gamma_0\right]} \tag{15}$$

Proof: Due to the page limit, we skip the proof here.

A. Max-Min with Self-interference Relay Selection: It is possible to calculate the corresponding outage probability for the high SNR regime as

$$F_{MMSI}^{SPC}(\gamma_0) = \left[P_{out}^{SPC}(\gamma_0)\right]^N, \tag{16}$$

as already mentioned in (15).

B. Max-Min Relay Selection: We can have end-to-end SNR of HD DF relaying scheme as

$$\begin{cases} \gamma_{R_i} = P_S\gamma_{SR_i} \\ \gamma_D = P_R\gamma_{R_iD} \end{cases} \tag{17}$$

The best relay selection scheme yields, higher performance is obtained when the signal transmit through N relays as

$$F_{MMRS}^{SPC}(\gamma_0) = \left[P_{out}^{SPC_HD}(\gamma_0)\right]^N \tag{18}$$

where $P_{out}^{SPC_HD}(\gamma_0) = 1 - e^{-\left(\frac{1}{\bar{\gamma}_{SR}P_S} + \frac{1}{\bar{\gamma}_{RD}P_R}\right)\gamma_0}$.
Proof: Due to the page limit, we skip the proof here.

II. Direct link (Non-EH without relaying system): End-to-end SNR from source to destination is depicted as

$$\gamma_{SD} = \frac{P_S|h_{SD}|^2}{\sigma_D^2}. \tag{19}$$

Outage probability is calculated by

$$P_{out}^{DR} = Pr(\gamma_{SD} < \gamma_0) = 1 - e^{-\frac{\gamma_0\sigma_D^2}{P_S\bar{\gamma}_{SD}}}. \tag{20}$$

5 Numerical

In this section, we provide some numerical examples for outage probability performance of the proposed multiple relay selection schemes. Unless otherwise stated, the source transmission rate is set $R_0 = 2$ bps/Hz and the energy harvesting efficiency is set to $\eta = 1$. Additionally, the simulation assumes a normalized system bandwidth. For simplicity, the similar noise variances are supposed at the relay (R) and destination (D) nodes, $\sigma_D^2 = \sigma_R^2 = \sigma^2 = 2 \times 10^{-2}$.

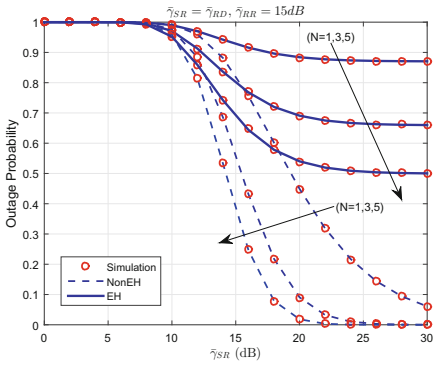


Fig. 2. Exact and simulation results of outage probability for MMSI vs. $\bar{\gamma}_{SR}$

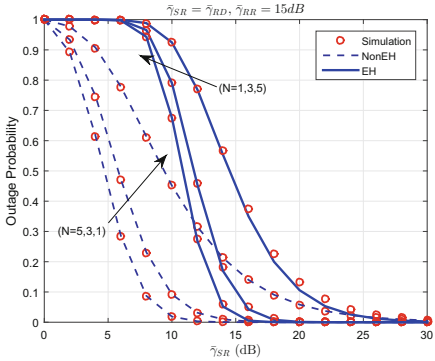


Fig. 3. Exact and simulation results of outage probability for MMSR vs. $\bar{\gamma}_{SR}$

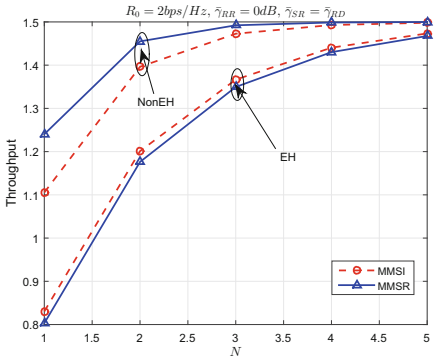


Fig. 4. The results of throughput for MMSR and MMSI vs. N relays with $\bar{\gamma}_{SR} = 0\,\mathrm{dB}$

In Figs. 2 and 3, the simulation and analytical results for outage probability are presented under various values of $\bar{\gamma}_{SR}$, respectively. We observe two figures that the simulation results match well with analytical results. It is apparent that

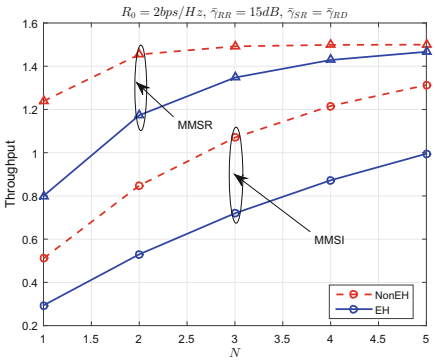


Fig. 5. The results of throughput for MMSR and MMSI vs. N relays with $\bar{\gamma}_{RR} = 20\,\text{dB}$

there is an downward trend in the results of outage probability in both MMSI and MMSR situations. We can see that the outage probability of EH is greater than the outage probability of Non-EH in both MMSI and MMSR.

In Figs. 4 and 5, we depict the throughput of MMSI and MMSR with N relays when we change the impact of self-interference ($\bar{\gamma}_{RR} = 0\,\text{dB}$ and $\bar{\gamma}_{RR} = 20\,\text{dB}$), respectively. It is clear that is an increase in both MMSI and MMSR instances. When the value of self-interference moves from $\bar{\gamma}_{RR} = 0\,\text{dB}$ to $\bar{\gamma}_{RR} = 15\,\text{dB}$, the results of throughput decrease remarkably, so it proves that self-interference cancellation need be resolved for improving performance. It is worth noting that the values of throughput of MMSI are nearly the same the results of MMSR when $\bar{\gamma}_{RR} = 0\,\text{dB}$, because when the self-interference is cancelled, the MMSI will be become MMSR.

6 Conclusion

In this study, we have developed opportunistic multiple relay selection schemes with both FD and HD scheme for both EH and Non-EH relaying mode in DF transmission mode. The derivative of expressions of outage probability for the proposed protocol is presented in this paper. The theoretical analysis was perceived to match the corresponding in the numerical results well. Furthermore, numerical results show that some parameters, including N which is the number of relays, the residual self-interference etc., all greatly affect the presence of multiple relays greatly affect the systems performance.

Appendix A

Proof of EH MMSI:
PDF and CDF of related variables as PDF of the SNR for self-interference

$$f_{\gamma_{RiRi}}(\gamma_0) = \frac{1}{\bar{\gamma}_{RR}} e^{\left(-\frac{\gamma_0}{\bar{\gamma}_{RR}}\right)}, \tag{A.1}$$

where the average residual self-interference power is denoted by $\bar{\gamma}_{RR}$. Thus, the CDF of the SNR at relay

$$F_{\gamma_{Ri}}(\gamma_0) = 1 - \left(1 - e^{\left(-\frac{(1-\beta)}{\gamma_0 \rho \bar{\gamma}_{RR}}\right)}\right). \tag{A.2}$$

The outage probability is given by

$$\begin{aligned} F_{\gamma_D}(\gamma_0) &= Pr\left(\rho \gamma_{SRi} \gamma_D < \gamma_0\right) = Pr\left(\gamma_D < \frac{\gamma_0}{\rho \gamma_{SRi}}\right) \\ &= \int_{x=0}^{\infty} f_{\gamma_{SRi}}(x)\left(1 - e^{\left(-\frac{\gamma_0}{x \rho \bar{\gamma}_{SR}}\right)}\right) dx \\ &= 1 - \sqrt{\frac{4\gamma_0}{\rho \bar{\gamma}_{SR} \bar{\gamma}_D}} K_1\left(\sqrt{\frac{4\gamma_0}{\rho \bar{\gamma}_{SR} \bar{\gamma}_D}}\right) \end{aligned}, \tag{A.3}$$

where $f_{\gamma_{SRi}}(x) = \frac{1}{\bar{\gamma}_{SR}} exp\left(-\frac{x}{\bar{\gamma}_{SR}}\right)$, with $x \geq 0$ and $K_1(.)$ is the first-order modified Bessel function of the second kind. Additionally, the end-to-end SNR is $z = \min\{\gamma_{Ri}, \gamma_D\}$. Then, the outage probability of overall system is

$$F_z(\gamma_0) = 1 - [1 - F_{\gamma_{Ri}}(\gamma_0)] \times [1 - F_{\gamma_D}(\gamma_0)] \tag{A.4}$$

We have the overall outage probability

$$P_{out}^{SPC_FD}(\gamma_0) = 1 - \left(1 - e^{\left(-\frac{(1-\beta)}{\gamma_0 \rho \bar{\gamma}_{RR}}\right)}\right) \sqrt{\frac{4\gamma_0}{\rho \bar{\gamma}_{SR} \bar{\gamma}_D}} K_1\left(\sqrt{\frac{4\gamma_0}{\rho \bar{\gamma}_{SR} \bar{\gamma}_D}}\right). \tag{A.5}$$

This ends the proof Appendix A.

Appendix B

Proof of EH MMSR:
The CDF of SNR in link S-R as

$$F_{\gamma_{Ri}}(\gamma_0) = 1 - \left(1 - e^{-\left(\frac{\gamma_0}{\bar{\gamma}_{SR} P_S (1-\beta)}\right)}\right). \tag{B.1}$$

The PDF of random variances γ_D is depicted in (A.3) as

$$F_{\gamma_D}(\gamma_0) = 1 - \sqrt{\frac{4\gamma_0}{\rho \bar{\gamma}_{RD} \bar{\gamma}_{SR}}} K_1\left(\sqrt{\frac{4\gamma_0}{\rho \bar{\gamma}_{RD} \bar{\gamma}_{SR}}}\right) \tag{B.2}$$

In DF scheme, the outage probability is given by

$$\begin{aligned} F_z(\gamma_0) &= 1 - [1 - F_{\gamma_{Ri}}(\gamma_0)] \times [1 - F_{\gamma_D}(\gamma_0)] \\ &= 1 - \left(1 - e^{-\left(\frac{\gamma_0}{\bar{\gamma}_{SR} P_S (1-\beta)}\right)}\right) \cdot \sqrt{\frac{4\gamma_0}{\rho \bar{\gamma}_{RD} \bar{\gamma}_{SR}}} K_1\left(\sqrt{\frac{4\gamma_0}{\rho \bar{\gamma}_{RD} \bar{\gamma}_{SR}}}\right) \end{aligned} \tag{B.3}$$

This ends the proof.

References

1. Niyato, D., Wang, P.: Delay-limited communications of mobile node with wireless energy harvesting: performance analysis and optimization. IEEE Trans. Veh. Technol. **63**, 1870–1885 (2014)
2. Luo, Y., Zhang, J., Letaief, K.B.: Relay selection for energy harvesting cooperative communication systems. In: Proceedings of the IEEE Global Communications Conference (GLOBECOM), pp. 2514–2519 (2013)
3. Ding, Z., Krikidis, I., Sharif, B., Poor, H.V.: Impact of channel state information on wireless energy harvesting cooperative networks with spatially random relays. In: Proceedings of the IEEE International Conference on Communications (ICC), pp. 4072–4076 (2014)
4. Nasir, A.A., Zhou, X., Durrani, S., Kennedy, R.A.: Relaying protocols for wireless energy harvesting and information processing. IEEE Trans. Wireless Commun. **12**, 3622–3636 (2013)
5. Yanju, G., Aissa, S.: RF-based energy harvesting in decode-and-forward relaying systems: ergodic and outage capacities. IEEE Trans. Wireless Commun. **14**, 6425–6434 (2015)
6. Xiao, Y., Han, Z., Da Silva, L.A.: Opportunistic relay selection for cooperative energy harvesting communication networks. In: Proceedings of the IEEE Global Communications Conference, pp. 4371–4376 (2014)
7. El Azzouni, S., Ercetin, O., El-Keyi, A., El Batt, T., Nafie, M.: Full-duplex cooperative cognitive radio networks. In: Proceedings of the 13th International Symposium on Modeling and Optimization in Mobile, Ad Hoc, and Wireless Networks (WiOpt), pp. 475–482 (2015)
8. Zhuo, C., Peng, X., Zhiguo, D., Xuchu, D.: The application of SWIPT to a cooperative full duplex network. In: International Symposium on Wireless Communication Systems (ISWCS), pp. 426–430 (2015)
9. Wu, T., Yang, H.C.: RF energy harvesting with cooperative beam selection for wireless sensors. IEEE Wireless Commun. Lett. **3**, 585–588 (2014)
10. Dinh-Thuan, D., Dinh-Thanh, V.: Performance of subspace based semi-blind channel estimation in MIMO systems. In: Proceedings of the International Conference on Networking and Information Technology, pp. 231–234 (2010)
11. Cheng, Z., Devroye, N., Liu, T.: The degrees of freedom of full-duplex bidirectional interference networks with and without a MIMO relay. IEEE Trans. Wireless Commun. **15**, 2912–2924 (2016)
12. Krikidis, I., Suraweera, H.A., Smith, P.J., Yuen, C.: Full-duplex relay selection for amplify-and-forward cooperative networks. IEEE Trans. Wireless Commun. **11**, 4381–4393 (2012)
13. Wang, Y., Xu, Y., Li, N., Xie, W., Xu, K., Xia, X.: Relay selection of full-duplex decode-and-forward relaying over Nakagami-m fading channels. IET Commun. **10**, 170–179 (2016)
14. Moualeu, J.M., Hamouda, W., Hongjun, X., Takawira, F.: Power assignment in multi-relay adaptive DF cooperative networks. In: Proceedings of the IEEE Global Communications Conference (GLOBECOM), pp. 2444–2449 (2012)
15. Upadhyay, M.A., Kothari, D.K.: Optimal resource allocation techniques for cooperative AF and DF wireless networks. In: Proceedings of the International Conference on Signal Processing and Integrated Networks (SPIN), pp. 382–387 (2014)

16. Zhou, X., Zhang, R., Ho, C.K.: Wireless information and power transfer: architecture design and rate-energy tradeoff. IEEE Trans. Commun. **61**(11), 4754–4767 (2012)
17. Bletsas, A., Khisti, A., Reed, D., Lippman, A.: A simple cooperative diversity method based on network path selection. IEEE J. Sel. Areas Commun. **24**, 659–672 (2006)

Pattern Discovery in the Financial Time Series Based on Local Trend

Mai Van Hoan[1(✉)], Dao The Huy[1], and Luong Chi Mai[2]

[1] University of Information and Communication Technology, Thai Nguyen University, Thai Nguyen, Vietnam
{maihoan,dthuy}@ictu.edu.vn
[2] Institute of Information Technology, Vietnamese Academiy of Science and Technology, Hanoi, Vietnam
lcmai@ioit.ac.vn

Abstract. We introduce a method for discovering new patterns in financial time series. Our method focuses on two main tasks of time series mining: Start with time series representation which helps to reduce the dimension and extracts useful feature of raw time series; Next discover on symbolic time series to find out new useful patterns which helpfully to improve trading decision in financial domain. In our work, we are interested in some patterns which have high win ratio percent (i.e. greater 70 %). In the first phrase, (i) raw data will be split into some segments with same length, (ii) local trend will be used to convert each subsequence into symbolic (U, u, s, d or D). In second phrase, we use a sliding window with size w moved on symbolic time series to create a collection of transactions. Based on this collection, the SPAM algorithm is used to discover all patterns with low *minSup*. In the last phrase, win/loss constraint used to discover new patterns in financial time series will be presented. Our demonstrate based on Gold Spot dataset from 2012-01-01 to 2015-01-01 is experimented.

Keywords: Local extraction · Time series mining · Representation · Pattern discovery · Evaluation · Meta trader 4 · Gold Spot

1 Introduction

Time series data is collection of interested real value which is recorded at regular time intervals [2]. Time series is importance in a large of fields, such as weather forecasting, server fault monitoring, financial forecasting, etc. Time series mining attracts many researchers to understand the time series data and applying it to solve real problems. Let us explain how time series information can improve business decision. As you known, the secret of Japanese which helped them leading the trend of rice market was revealed the middle 1990s by Steve Nison. The most important question is: How could they do like that? And the answer was revealed "Candlestick". They used the price rice represented into Japanese Candlestick, discovered special Candlestick patterns (Fig. 1A) and trading rice

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_48

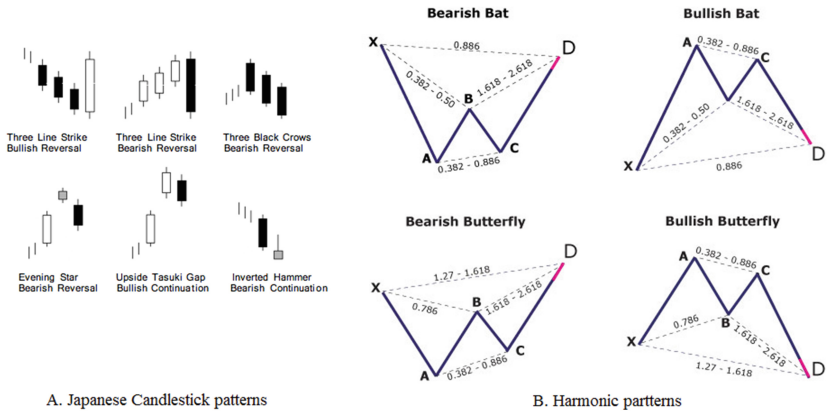


Fig. 1. Japaneses Candlestick and Harmonic patterns - Source: Internet

decision based on its. And recent today, Hamonic (Fig. 1B) is a new pattern in the financial domain. In this article, we present a method to find out new "Symbolic patterns" which helpfully to improve trading decision. A symbolic pattern is a collection of symbol (i.e. AABBDCCD - Bullish, CDEDS - Bearish). Our method provides a complete solution starts with time series preprocess, then time series symbolization, after that new patterns discovery and finally evaluating pattern with win/loss constraint. The rest of this paper is organized as follows: Sect. 2 introduces time series definition, several methods of time series representation, and Foreign exchange market; Sect. 3 provides our methodology used to discover new patterns in the financial time series; In Sect. 4, we present experiment of our method; finally, in Sect. 5 it shows our conclusions and perspective.

2 State of the Art

The following section, we present terminology, definition and methodology. We start with introduction to time series data:

2.1 Time Series

Time series: A time series $T = \{t_1, t_2, ...t_m\}$ knows as a sequence of values at regular time intervals [3]. *Subsequence :* A subsequence C_p of T is a time series of length w that is extracted from the time series T of length m, starts at position p and finishes at position $(p + w - 1)$ from T [3].

$$Cp = \{t_p, ..., t_{p+w-1}\}, p \in [1, .., m - w + 1] \tag{1}$$

Sliding window : A window is a slice of time of width w, where w can be expressed either in terms of a real time duration or in terms of a number of consecutive measures. A sliding window is a window that is moved along a time series T to create all possible subsequences of T of length w.

2.2 Time Series Representation

Time series representation is one of popular task in time series mining, this task always uses to reduce the dimensionality of raw time series. In this section, we focus on two methods presented by E. Keogh and J. Lin.

PAA - Piecewise Aggregate Approximation: E. Keogh *et al.* presented a method reducing dimension of time series called PAA - Piecewise Aggregate Approximation [1] since 2001. PAA has several advantages: fast to compute, arbitrary length queries, etc. This method reduce a time series from n - high dimensions to N - low dimensions by split original time series into N segments of same length.

SAX - Symbolic Aggregate approXimation: SAX- Symbolic Aggregate approximation [5] is presented the first time by J. Lin (2003). SAX is a symbolic representation of time series, which allows a time series of length n to be reduced to a string of length w ($w \ll n$).

2.3 Forex -Foreign Exchange Market

The foreign exchange market, also acknowledged as Forex or Fx for short, is a currency trading market spread all around the globe. Central banks, large financial institutions, commercial banks, hedge funds, and extremely wealthy individuals for currency trading traditionally use the Forex market. However, by the appearance of the Internet and its augment, the Fx market became available for small retailers [7]. According to Hossein Talebi et al. [7] In Forex market trading is done by selling and buying currency pairs, i.e. EUR/USD. There are several currency pairs, although the major ones in term of the amount of daily transactions are Euro vs US Dollar (EUR/USD), Australian Dollar vs US Dollar (AUD/USD), Great Britain Pound vs US Dollar (GBP/USD), US Dollar vs Canadian Dollar (USD/CAD), US Dollar vs Swiss Franc (USD/CHF) and US Dollar vs Japanese Yen (USD/JPY). According to the Bank for International Settlements [9], in the Contract For Differences – CFD market besides the currency trading there are other tradable instruments such as indices like: DJI, SP500, Nd100... energy as Oil, Gas... and metals example spot gold, spot silver... In the Bank for International Settlements [8] showed that average of daily exchange in foreign exchange markets is $5.3 trillion in April 2013. And Hossein Talebi et al. [7] wrote that huge amount of turnover makes Forex market the largest trading market in the world which is approximately 160 times larger than the New York Stock Exchange.

3 Our Approach

In this section, we propose a method to discover new patterns in financial time series (Forex in particular). As show in Fig. 2, our method will be performed in three phrases: Symbolic time series representation; Pattern discovery and pattern

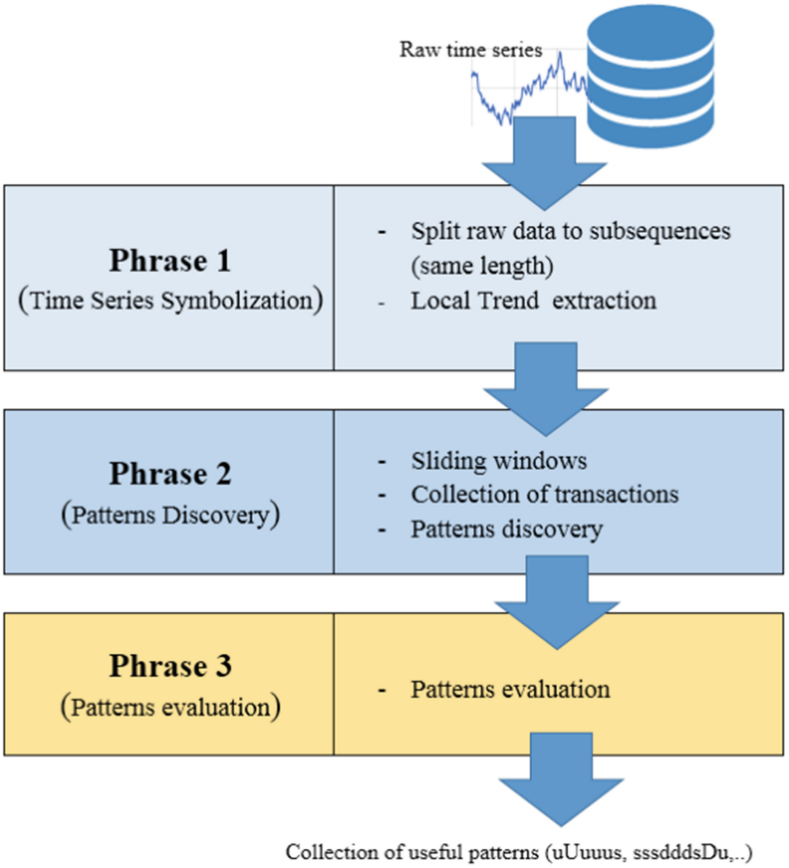


Fig. 2. Pattern discovery in the financial time series based on local trend

evaluation. In the first phrase, we propose a method to convert a raw time series into symbolic time series based on local trend extraction and cycle characteristic of time series. In second phrase, A collection of pattern will be discovered by using a slide window and sequential patterns mining. The last phrase, we show the method to evaluate each pattern with win/loss constraint.

Let us define some notations are used in our method:

- A raw time series $T = \{t_1, t_2, ..., t_m\}$ where m is length of time series
- Number point of each segment: p
- A set of symbol $S = U, u, s, d, D$ where U standards for strong trend; u: sideway up; s : sideway; d : sideway down; D : strong down.
- A symbolic time series of T: $\bar{T} = \{\bar{t_1}, ..., \bar{t_k}\}$ where $k = \frac{m}{p}$
- A sliding window of size w
- A database of transactions D
- A support threshold of itemset frequent: $minSup$

– A collection of pattern C, which created from D with support threshold $minSup$
– A threshold of winning trade $perWin$
– A collection of new pattern in the financial time series $\bar{C}$ which created from C with win/loss constraint.

3.1 Time Series Symbolization Based on Local Trend Extraction

In this section, we present a method to transform a raw time series into symbolic time series. The object of our method to reduce noise and dimension of original time series.

Local financial trend defination in time series: In financial time series, we have: *Up trend :* Financial asset, market is in up trend when its price movement is rising in overall direction. It has each successive peak and trough is higher than the ones found earlier in the trend; *Down trend :* Financial asset, Bull market is in down trend when its price movement is sinking in overall direction. It has each successive peak and trough is lower than the ones found earlier in the trend; *Sideway :* A market is in sideway status when it has price within a range. They don't make higher highs (lower lows) or a breakout above the previous peak (trough) price. ***Local trend***: Trend was extracted from interval time of time series.

Our method consist of three steps:

Step 1: Time series segmentation: Original time series will be segmented into subtimeseries with same length based on its cyclic characteristic (exp: each 15 min, 30 min, hour, 4 h,... etc.) using the Algorithm 1.

Algorithm 1. Algorithm to segment

Require: A raw time series T with size m, number point of each segment p
Ensure: A collection of subsequence $\hat{T}$

```
T̂ = ∅
for j ← 1, n do
    T̂_j ← {t_j, t_{j+1}, ..., t_{j+p}}
    T̂ ← T̂, {T̂_j}
    j ← j + p
return T̂
```

Step 2: Data normalization: To make comparison two sub-time series fairly, each segment will be normalize between 0 and 1 with uptrend and –1 and 0 with down trend. Algorithm 2 will used to normalize raw time series.

Algorithm 2. Algorithm to normalize

Require: A collection $\hat{T}$ with k subsequences
Ensure: A collection of subsequence $\hat{T}$ normalized
$maxDown \leftarrow 0$
$maxUp \leftarrow 0$
for $j \leftarrow 1, k$ **do**
  $sub \leftarrow \hat{T}_j$
  $change \leftarrow (sub_p - sub_0)$
  if $((change > 0) and (maxUp > change))$ **then** $maxUp \leftarrow change$
  if $((change < 0) and (maxDown < change))$ **then** $maxDown \leftarrow change$
for $j \leftarrow 1, k$ **do**
  $sub \leftarrow \hat{T}_j$
  $change \leftarrow (sub_p - sub_0)$
  if $(change > 0)$ **then**
    for $i \leftarrow 1, p$ **do**
      $sub_i \leftarrow (sub_i / maxUp)$
  if $(change < 0)$ **then**
    for $i \leftarrow 1, p$ **do**
      $sub_i \leftarrow (sub_i / maxDown)$
  $\hat{T}_j \leftarrow sub$
.
return $\hat{T}$

Step 3: Mapping Symbol: This is important step, which helps to assign each subsequence into a label U, u, s, d, D by using $threshold1$, $threshold2$ and local trend extraction. Each subsequence will be converted directly to symbolic by using local trend extraction with Algorithm 3.

3.2 Pattern Discovery with Support Threshold

We create a collection of transactions D by using a sliding window across symbolic time series $\bar{T}$.

Algorithm 4. Algorithm to create a collection of transaction

Require: A symbolic time series $\bar{T}$, a sliding window of size w
Ensure: A dataset of transactions D
$D = \emptyset$
for $i \leftarrow 1, sizeof(\bar{T})$ **do**
  $t_i \leftarrow$ subsequence of size w from $\bar{T}$
  $D \leftarrow t_i$
.
return D

Based on the collection D and a support threshold $minSup$, we find out all frequent patterns in the collection D by using SPAM algorithm. In general, we

Algorithm 3. Algorithm to symbolic time series

Require: A collection $\hat{T}$ with k subsequence normalized, $threshhold1$, $threshold2$ with $0 < threshold1 < threshold2$
Ensure: A symbolic time series $\bar{T}$
for $j \leftarrow 1, k$ **do**
$sub \leftarrow \hat{T}_j$
$change \leftarrow (sub_p - sub_0)$
if $(change > threshhold2)$ **then**
$\bar{T}_j \leftarrow' U'$
if $(change > threshold1) and (change < threshhold2)$ **then**
$\bar{T}_j \leftarrow' u'$
if $(change < (-1) * threshold1) and (change < threshhold1)$ **then**
$\bar{T}_j \leftarrow' s'$
if $(change < (-1) * threshold1) and (change > (-1) * threshhold2)$ **then**
$\bar{T}_j \leftarrow' d'$
if $(change < (-1) * threshhold2)$ **then**
$\bar{T}_j \leftarrow' D'$
.
return $\bar{T}$

have some good patterns and some bad patterns. In next step, we will present the way to collect useful patterns with win/loss constraint.

3.3 Pattern Evaluation with Win/Loss Constraint

In Financial Time Series, some types of patterns like Japanese Candle pattern, Harmonic pattern help to make good decision in market. One of the most import constraint that is win/loss ratio. In situation, each pattern has win ratio greater than $perWin$ (i.e. 50 %, 65 %, 70 % win rate, etc.) is **useful pattern**. In this step, We evaluate each patterns in C by using simulation in interval time and obtain useful pattern with $perWin$ ratio.

4 Experimentation

4.1 Dataset of Financial Time Series

In our experimentation, we use Gold Spot close price from 2012-01-01 to 2015-01-01 in Forex market. The data was recorded in minute interval.

4.2 Framework

SPMF framework and Metatrader 4 was used in our case. SPMF - Sequential Pattern Mining Framework is introduced by Philip Fournier-Viger, the SMPF is a Java open source. It is a powerful framework in pattern mining (itemset mining, sequential pattern mining, periodic pattern mining, etc.), easy to use

with command line or user interface. It's content of 120 data mining algorithms (June 2016).

MetaTrader 4 - MT4 is a trading platform used by both of trader and broker. MT4 contents some functions: Trade Execution, live Forex chart price tick by tick, ...etc. and a good function which help to evaluate our patterns that is **back testing** with historical data.

4.3 Symbolic Representation of Financial Time Series

Our raw Gold Spot time series T of length 1.456.080 values ($m = 1.456.080$). We split raw data into subsequences with $p = 60$ (i.e. one hour) using Algorithm 1. The Algorithm 2 used to normalize each subsequence. After we transform raw dataset to symbolic time series by using Algorithm 3, we use Algorithm 4 to create collection of transaction with sliding windows size 13.

4.4 New Pattern Discovery

Based on the collection of transaction in bellow steps, we use SPAM algorithm with small support threshold $minSup$ (i.e. 1 %, 5 %) to discover maximal pattern which hidden inside collection D. New pattern will be discover by using MT4 back testing simulation based on win/loss constraint (i.e. $perWin$ ratio: 50 %, 65 %, 70 %);

4.5 Results

The result of our experimentation is presented in Figs. 3 and 4. Figure 3 is a piece our symbolic time series which transform from raw data by using our algorithm. Figure 4 (A) presents the symbols distributed in our symbolic time series, the result is very similar nature of Forex Market (70 %–80 % market conditional is

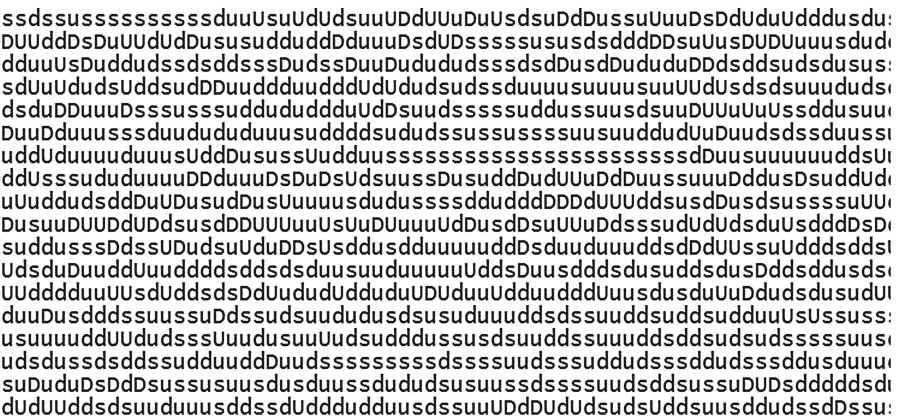
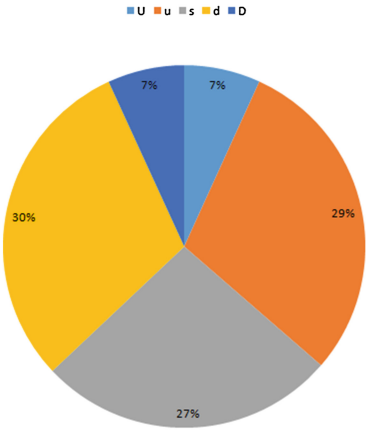


Fig. 3. A piece of symbolic time series



(A)

sd	sduU	su	dDU	DU
sdD	sdU	suU	du	uD
sdDu	sD	sU	duU	
sdDuD	sDu	du	dU	
sdDU	sDuU	dDu	Du	
sdu	sDU	dDuU	DuU	

(B)

BUY	SELL
sdDuU	sd
sdDuU	sduU
sDuU	
dDuU	

(C)

Fig. 4. (A) Symbols distributed; (B) Collection of pattern; (C) Useful patterns

sideway market). Figure 4 (B) show all pattern which discovered by using SPAM algorithm with *minSup* from 1 %. The propose of small *minSup* which helps to find maximal pattern candidate. Figure 4 (C) gives you high recommended patterns which helps to make good decision with win percent greater than 70 % (its means when a pattern is appear in market, we enter an Order and Exit in next one hour).

5 Conclusions

In this paper we presented a method to discover new useful pattern in the financial time series. Our method quickly find out patterns and easily evaluate with win/loss constraint. New patterns found by using our method could be used directly in making trading decision in the real market.

Our method consists of three phrases: time series symbolization based on local extraction, pattern discovery and pattern evaluation using win/loss constraints. First, raw time series data will be divided into subsequences and "local trend extraction" used to convert each segment into symbolic time series. Second, a transaction database will be created based on symbolic time series by using a sliding window. Then, all pattern will be discovered by using SPAM algorithm with low *minSup*. Finally, new pattern is determined by using win/loss constraint. Our experimentation based on collection of GOLD Spot.

In further work, we will change another constraint (i.e. Risk-Reward constraint) to discover pattern in the financial time series. And using different methods to convert original time series into symbolic time series like: SAX, time series symbolization based on clustering.

References

1. Keogh, E., Chakrabarti, K., Pazzani, M., Mehrotra, S.: Dimensionality reduction for fast similarity search in large time series databases. Knowl. Inf. Syst. **3**(3), 263–286 (2001)
2. Eschenfelder, A.H.: Data Mining and Knowledge Discovery Handbook. Solid-State Sciences. Springer, Heidelberg (2010)
3. Keogh, E., Lin, J., Truppel, W.: Clustering of time-series subsequences is meaningless: implications for previous and future research. Knowl. Inf. Syst. **8**(2), 115–122 (2003)
4. Abraham, B., Ledolter, J.: Monthly sales of U.S. houses (thousands) 1965–1975 (1983)
5. Lin, J., Keogh, E., Lonardi, S., Chiu, B.: A symbolic representation of time series, with implications for streaming algorithms. In: Proceedings of the 8th ACM SIGMOD Workshop on Research Issues in Data Mining and Knowledge Discovery DMKD 2003, (16) p. 2 (2003)
6. Ayres, J., Flannick, J., Gehrke, J., Yiu, T.: Sequential pattern mining using a bitmap representation. In: Proceedings of the Eighth ACM SIGKDD International Conference on Knowledge Discovery and Data Mining KDD 2002, p. 429 (2002)
7. Talebi, H., Hoang, W., Gavrilova, M.L.: Multi-scale foreign exchange rates ensemble for classification of trends in forex market. Procedia Comput. Sci. **29**, 2065–2075 (2014). 14th International Conference on Computational Science ICCS 2014
8. 2016 Triennial Central Bank Survey of Foreign Exchange and OTC Derivatives Markets. http://www.bis.org/press/p160330.htm
9. OTC derivatives statistics at end-June 2013. http://www.bis.org/publ/otc_hy1311.pdf

Performance Evaluation of SAR ADC with Organic Semiconductor

Huyen Thanh Pham[1,2(✉)], Toan Thanh Dao[2], and Thang Vu Nguyen[1]

[1] Hanoi University of Science and Technology, No.1, Dai Co Viet, Hai Ba Trung, Hanoi, Vietnam
[2] University of Transport and Communications, No.3, Cau Giay, Lang Thuong, Dong Da, Hanoi, Vietnam
huyenktdt@utc.edu.vn

Abstract. This paper presents impact of sampling and input signal frequencies on dynamic performance of a fully differential organic 6 bit SAR ADC. The simulation results proved that the total power consumption of the circuit increased linearly with the sampling frequency and was almost stable in the wide range of the input signal frequency. At the sampling frequency of 1 KHz and the input signal frequency of 10 Hz, the power dissipation, *ENOB*, *SFDR*, *THD*, and *FoM* are 443.1 μW, 4.83 bit, 37.71 dB, −34.01 dB, and 15.60 nJ/conv, respectively.

Keywords: Analog-to-digital converter · Organic integrated circuit · Dynamic performance · Modeling and simulation · Low-power consumption

1 Introduction

An analog-to-digital converter (ADC) is considered as one of the most important parts in electronic systems since it influences directly on the accuracy and stability of digitalization. Whereas, performance of an ADC depends on its input parameters, such as sampling frequency and input signal frequency. Silicon-based ADCs' specifications have been analyzed widely in many researches [1,2], but there is no reported one for organic ADCs, which have different characteristics from their counterparts. Indeed, organic integrated circuits recently have become attractive due to their outstanding advantages, for example, flexibility, bio-compatibility, and low-cost manufacturing [3–6]. They, however, have suffered several obstacles, such as stability, age degrading, low-speed, and large-power consumption [4,7,8]. Those issues have revealed that organic ADCs should be evaluated properly before applying them in practical systems.

In this research, the dynamic parameters including an effective number of bit (*ENOB*), a signal to noise and distortion ratio (*SINAD*), a spurious free dynamic range (*SFDR*), and a total harmonic distortion (*THD*) of an organic successive-approximation register (SAR) ADC were examined in various initial setup conditions. In addition, we would like to analyze the ADC in terms of

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_49

energy dissipation and quality; thus, total power consumption (P) and a figures-of-merit (FoM) would be calculated to find out the most appropriate set of input specifications.

The rest of the paper is organized as follows. The schematic design and simulation setup conditions are described in Sect. 2. In Sects. 3 and 4, we introduce how to define dynamic parameters and compare our simulated results to present state-of-the-art ones. Lastly, Sect. 5 concludes the study.

2 Schematic Design and Simulation Setup

The fully differential 6 bit SAR ADC were designed and simulated by using two types of newly modeled organic devices called pentacene P-channel and fullerene N-channel. The detailed descriptions of the organic field-effect transistor (OFET) fabrication, electrical characterization, modeling, and simulation processes were reported in our previous works [9–11]. In brief, the OFETs were fabricated on the SOI (Silicon-On-Insulator) wafer. Then, the physical and electrical characteristics of the transistors were measured. These experimental data were used to extract the main parameters and to probe the fitting ones declared in the organic transistor models. With the modeled elements, which were supplemented to the device library of the organic process design kit (OPDK) tool added in Cadence Virtuoso environment [12], we have designed a number of analog and digital circuits composing logic gates, flip-flops, comparators, and ADCs.

In our work, the SAR ADC was used as an object to observed impact of input parameters on. Figure 1 presents the simplified schematic of the ADC wherein main blocks consist of two bootstraps to track and hold input signals, a comparator to define output bits, a clock generator named clk-gen to create controlling clock pulses, two digital-to-analog converters (DACs) to regenerate output analog signals, and a register called Dout to store digital outputs. To reduce the power dissipation of the circuit, simple structures with clock pulses are chosen. For example, dynamic-latched comparators and capacitive switch DACs operate in synchronous mode and they have been known as most power saving ones [13–15]. Besides, the DAC circuit is halved the total capacitance comparing to that in the conventional N bit capacitive DAC because the most significant bit (MSB) is defined separately [15].

In all simulation runnings, the initial variables of the circuit were set at the resolution of 6 bit, the supply voltage of 10 V, and the reference voltage of 6.4 V.

To examine its effects on the ADC, the sampling frequency was increased from the Nyquist frequency of 20 Hz to 5 KHz while the input signal one was fixed at 10 Hz. The power consumption was taken into account when we compared the simulation results to find out the most appropriate sampling frequency used for later simulations. We set the sampling frequency at the found value and changed the input frequency to clarify which one would be the best choice in the same terms.

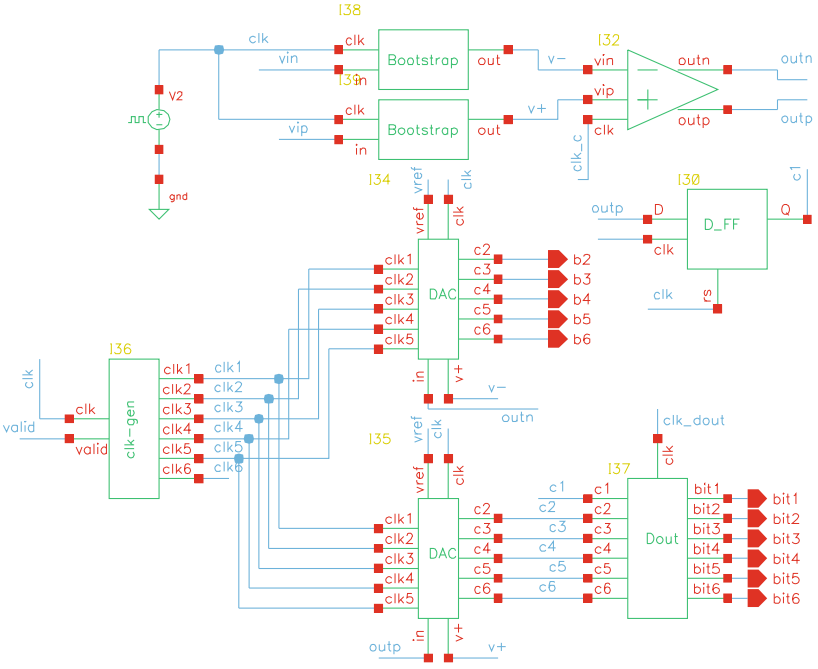


Fig. 1. Simplified schematic of the differential organic SAR ADC.

3 Dependence of Dynamic Performance on Sampling and Input Frequency

As above-mentioned, the *ENOB*, *SINAD*, *SFDR*, and *THD* are the four main ADC dynamic parameters supervised in relation to the power consumption P and the figures-of-merit *FoM* in order to evaluate performance of the ADC. At first, the digital output signal of the ADC was regenerated to analog formation in the time domain by an ideal DAC with the same resolution. Next, by using Fast-Fourier-Transform (FFT) method, the reproduced waveform was transformed into the spectrum in the frequency domain. Then, four key dynamic specifications including *ENOB*, *SINAD*, *SFDR*, and *THD* were calculated from the output spectrum.

In known theories [16,17], the *SINAD* provides information of noise and harmonic energy appearing in the frequency spectrum. Based on it, the *ENOB* is expressed as

$$ENOB(dB) = \frac{SINAD(dB) - 1.76}{6.02}. \tag{1}$$

Equation (1) shows that the *ENOB* is linear with the *SINAD*; therefore, plotting both of them is inessential. In this paper, the ENOB is plotted, and it is noted that the maximum value of *ENOB* is the default resolution of the ADC.

The *SFDR* is the ratio of the rms amplitude of the fundamental tone to the rms value of the largest distortion component in a specified frequency range. The *SFDR*, which indicates harmonic characteristics, becomes crucial since noise and harmonics limit an ADC's dynamic range. The higher the *SFDR* is, the better the ADC performance is [16].

As for distortion, the *THD* is defined as the rms summation of all harmonics in the output spectrum in which only first five harmonics are predominant. In general, the *THD* should be as low as possible [11].

Last but not least, the *FoM* is implied energy-per-bit versus the speed; hence it is widely used in literature for ADC performance comparison [1,2,17]. The *FoM* is expressed commonly as

$$FoM = \frac{P}{f_{\mathrm{S}} * 2^{\mathrm{ENOB}}}. \tag{2}$$

Where the f_S is the sampling frequency. The test-bench in Cadence was simulated in the transient mode with tens of cycles of the input signal to estimate the total power consumption of the circuit, and the P was the average value of the data. The Eq. (2) demonstrates that the *FoM* would be minimized if the f_S and/or *ENOB* is increased while the P is decreased.

4 Results and Discussion

Figure 2 presents the FFT spectra of the output signal at the different sampling frequency of 20, 100, 200, 500, 1000, and 2000 Hz and at the fixed input signal frequency of 10 Hz. Every bandwidth is equal to the half of corresponding f_S. Indeed, we also ran the simulations at the sampling frequency of 3 KHz and 5 KHz, but these attempts were collapsed. The results reveal that the circuit can not work at frequencies higher than 2 KHz. It could be due to the limited bandwidth of OFET devices [11].

Figure 3 represents the *ENOB*, *SFDR*, *P*, *THD* and *FoM* in the frequency domain. Figures 3(a,b) show that the P increases linearly with the f_S. The cause is that the transistor currents go up at the rising and falling edges of the clock pulses. This leads to larger power dissipation of the circuit when the f_S shifts to higher frequencies. Although the *SFDR* degrades at the frequency of 200 Hz, it resumes to the growing tendency. The *ENOB* and *SFDR* reach the maximum peaks of 5.09 bit and 38.80 dB at the f_S of 2 KHz. However, at this point, the total power consumption and *THD* are pushed to the highest values as well. That is inefficient in terms of energy saving.

As seen in Figs. 3(c,d), the *THD* and the *FoM* achieve lowest points of −34.01 dB and 15.60 nJ/conv respectively at the f_S of 1 KHz. The *ENOB* and *SFDR* also obtain highest peaks there. This proves that the optimal sampling frequency f_S is at 1 KHz because the ADC circuit consumes the least energy while maintaining the good *ENOB* of 4.83 bit and the high *SFDR* of 37.71 dB.

After choosing the optimal sampling frequency f_S, we set up the simulation to run over a wide range of input signal frequency f_{in}. The f_S was fixed at 1 KHz;

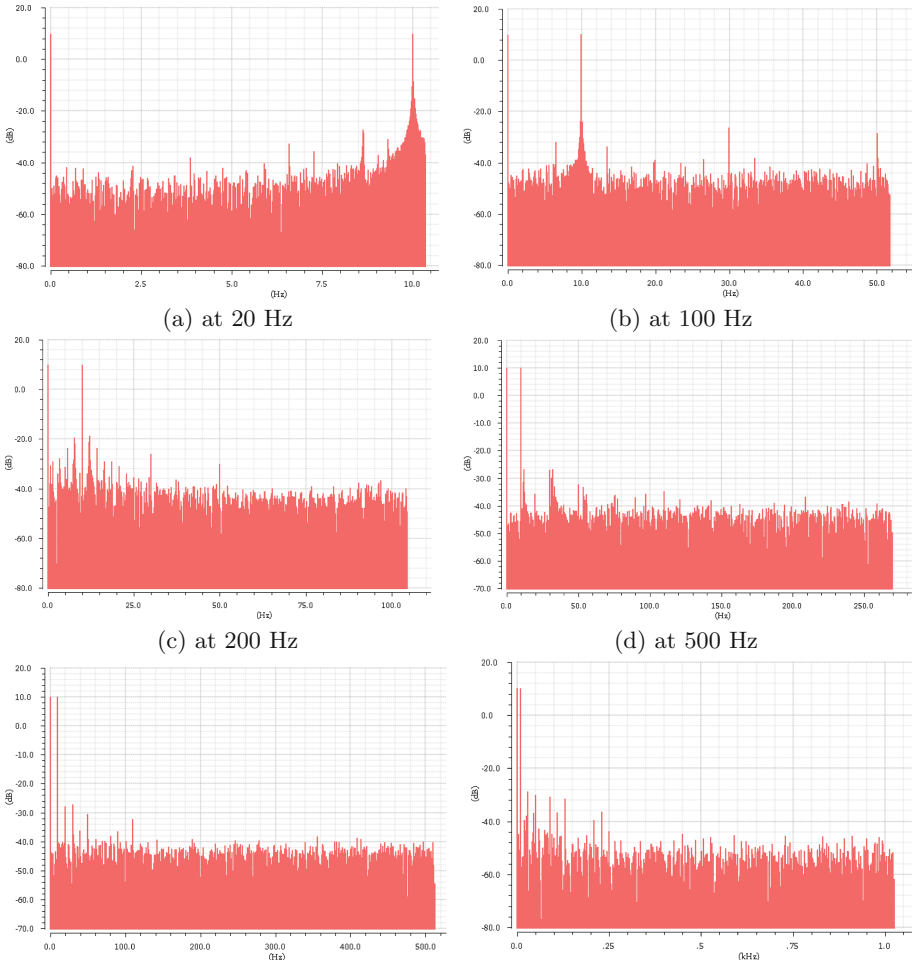


Fig. 2. Spectrum of output signal at different sampling frequency.

thus, the maximum f_{in} was at 500 Hz, as known from the Nyquist theorem. Here, we increased f_{in} from 1 to 500 Hz. Simulation results are analyzed and summarized in Fig. 4.

Figures 4(a,b) suggest that the P seems to be independent on the input signal frequency; it is almost stable at 440 μW. This results from the much dissipation of the circuit at the rising and falling edges of the clock pulses, which are fixed with the same f_S.

The *ENOB*, *SFDR*, and *FoM* change slightly at the input frequency from 1 to 10 Hz, as seen in Figs. 4(a,b,d). Nevertheless, the *ENOB* and *SFDR* rapidly decrease, not the *FoM* at the f_{in} higher than 10 Hz.

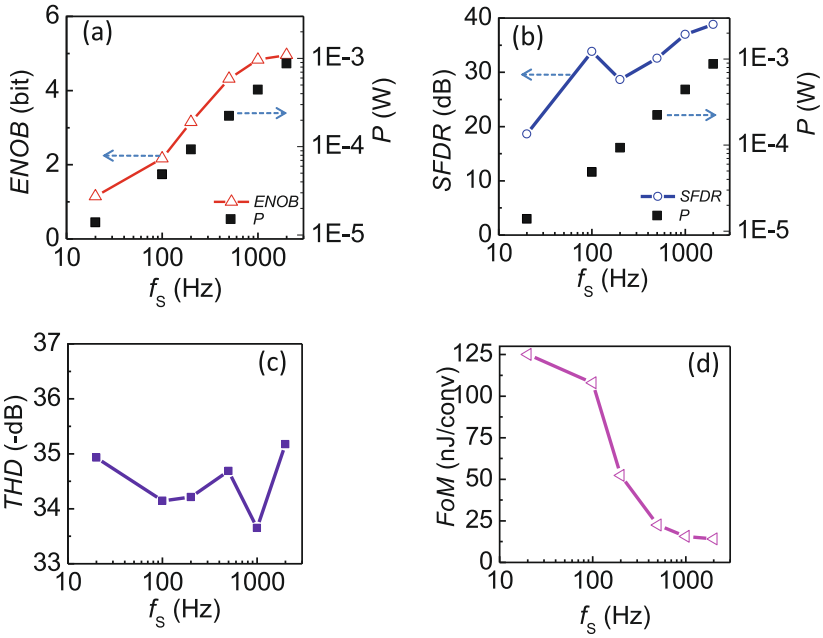


Fig. 3. Relationships of (a) *ENOB-f_S-P*, (b) *SFDR-f_S-P*, (c) *THD-f_S* and (d) *FoM-f_S*.

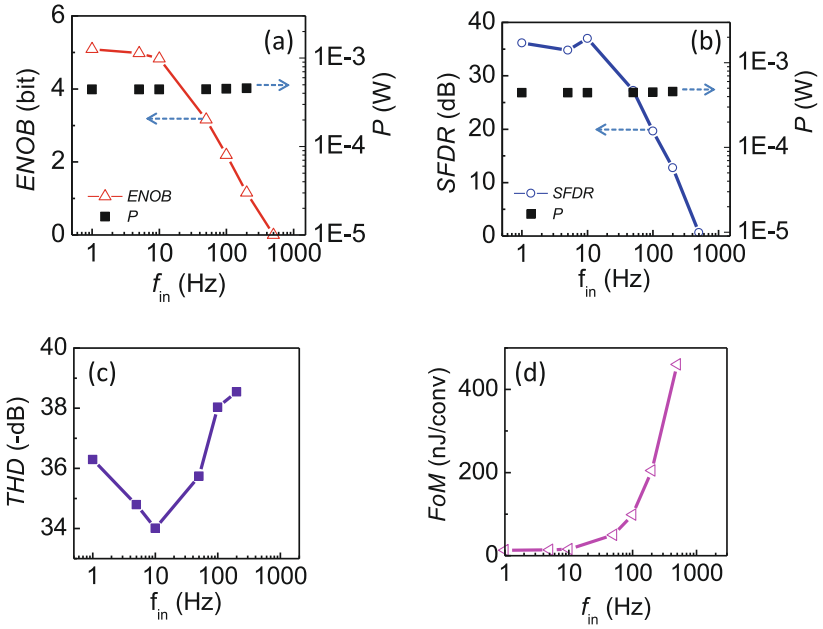


Fig. 4. Relationships of (a) *ENOB-f_{in}-P*, (b) *SFDR-f_{in}-P*, (c) *THD-f_{in}* and (d) *FoM-f_{in}*.

Table 1. Comparison of organic ADCs' performance

Research	Circuit structure	Supply voltage (V)	$ENOB$ (bit)	f_S (Hz)	$P(\mu W)$	FoM (nJ/conv)
[7]	Delta-Sigma	15	4.96	500	1500	116.22
[18]	VCO-based	20	6.00	40	480	187.50
[8]	SAR R-2R	40	2.96	4.17	540	16642.30
This work	C-SAR	10	4.83	1000	443.10	15.60

The *THD*, whereas, achieves the minimum peak of −34.01 dB at the f_{in} of 10 Hz, as presented in Fig. 4(c). This means, in effect, that the harmonic tones are much smaller than that in the other spectra. At the higher frequencies, the *THD* rises sharply, so it makes the *ENOB* fall rapidly, as pointed out in Fig. 4(a). Therefore, the optimal input signal frequency is chosen to be at 10 Hz.

Table 1 summarizes the comparison of our SAR ADC to other ADCs in [7,8,18]. Within the comparative parameters, the *FoM* is the most effective to evaluate an ADC's performance, and the *FoM* of our ADC working at the optimal input conditions is the smallest value of 15.60 nJ/conv. Besides, our proposed ADC dissipates the least power of 443.1 μW at the lowest supply voltage of 10 V and at the highest sampling frequency of 1 KHz.

5 Conclusions

We have estimated impact of the sampling frequency and the input signal frequency on dynamic performance of the organic 6 bit SAR ADC. The analyzed results show that at the same input signal frequency, the total power consumption rises linearly with the sampling frequency. However, at the same sampling frequency, the total circuit power consumption is approximately 440 μW with the various input signal frequency range. The optimal sampling and input signal frequencies are found to be 1 KHz and 10 Hz respectively at which the ADC obtains both high performance and low-power dissipation.

References

1. ADC Performance Survey 1997-2015. http://web.stanford.edu/~murmann/adcsurvey.html. (accessed 20 Aug 2016)
2. A/D-converter performance evolution (2012). https://converterpassion.wordpress.com/articles/a-d-converter-performance-evolution. (accessed 20 Aug 2016)
3. Sekitani, T., Someya, T.: Stretchable organic integrated circuits for large-area electronic skin surfaces. Mater. Res. Bull. **37**, 236–245 (2012)
4. Sekitani, T., Yokota, T., Someya, T.: Large-Area, Printed Organic Circuits for Ambient Electronics. Large Area and Flexible Electronics, 365–380 (2015)
5. Mandal, S., Noh, Y.: Printed organic thin-film transistor-based integrated circuits. Semicond. Sci. Technol. **30**, 064003 (2015)
6. Takeda, Y., Hayasaka, K., Shiwaku, R., Yokosawa, K., Shiba, T., Mamada, M., Tokito, S.: Fabrication of ultra-thin printed organic TFT CMOS logic circuits optimized for low-voltage wearable sensor applications. Sci. Rep. **6**, 25714 (2016)

7. Marien, H., Steyaert, M.S.J., Veenendaal, E., Heremans, P.: A fully integrated $\Delta\Sigma$ ADC in organic thin-film transistor technology on flexible plastic foil. IEEE J. Solid-State Circ. **46**, 276–284 (2011)
8. Abdinia, S., et al.: A 4-bit ADC manufactured in a fully-printed organic complementary technology including resistors. In: IEEE International Solid State Circuits Conference, pp. 106–107. IEEE Press, Washington, DC (2013)
9. Dao, T.T., Matsushima, T., Friedlein, R., Murata, H.: Controllable threshold voltage of a pentacene field-effect transistor based on a double-dielectric structure. Org. Electron. **14**, 2007–2013 (2013)
10. Pham, H.T., Nguyen, T.V., Sakai, H., Dao, T.T.: Design and analysis of tunable Vth pOTFT-based CMOS SRAM Cell. In: The 12th International Thin-Film Transistor Conference (ITC), Hsinchu, Taiwan, p. 70 (2016)
11. Pham, H.T., Nguyen, T.V., Pham-Nguyen, L., Sakai, H., Dao, T.T.: Design and simulation of a 6 bit successive-approximation ADC using modeled organic thin-film transistors. Act. Passive Electron. Compon. **2016** (2016)
12. The Organic Process Design Kit (OPDK), http://opdk.umn.edu/. (accessed 20 Aug 2016)
13. Jeon, H.J., Kim, Y.B.: A novel low-power, low-offset, and high-speed CMOS dynamic latched comparator. J. Analog Integr. Circ. Sig. Process. **70**, 337–346 (2012)
14. Yuan, C., Lam, Y.Y.H.: An ultra-low energy capacitive DAC array switching scheme for SAR ADC in biomedical applications. In: IEEE International Conference on IC Design & Technology (ICICDT), Kaohsiung, Taiwan, pp. 1–4 (2011)
15. Liu, C., Chang, S., Huang, G., Lin, Y.: A 10-bit 50-MS/s SAR ADC with a monotonic capacitor switching procedure. IEEE J. Solid-State Circ. **45**, 731–739 (2010)
16. Baker, R.J.: CMOS Circuit Design, Layout, and Simulation. John Wiley & Sons (2008)
17. Jonsson, B.E.: An empirical approach to finding energy efficient ADC architectures. In: IMEKO IWADC & IEEE ADC Forum, pp. 1–6 (2011)
18. Raiteri, D., et al.: An organic VCO-based ADC for quasi-static signals achieving 1LSB INL at 6b resolution. In: 2013 IEEE International Solid-State Circuits Conference, pp. 108–109 (2013)

Performance Evaluation of Wireless Networks Based on Testbed

Ngo Hai Anh[1(✉)], Takumi Tamura[2], and Pham Thanh Giang[1]

[1] Institute of Information Technology,
Vietnam Academy of Science and Technology, Hanoi, Vietnam
{ngohaianh,ptgiang}@ioit.ac.vn
[2] Department of Electrical Engineering,
Nagaoka University of Technology, Nagaoka, Japan
s121058@stn.nagaokaut.ac.jp

Abstract. Today the trend of using hardware environment for testing and assessmenting (testbed) is increasingly becoming more popular. This is obvious because testing on the real equipment is always more accurate than testing on the theoretical model or software simulation. In addition, the cost of computer hardware is decreasing quite fast, and its design is getting more compact, this contributes to facilitating the use of hardware system. Therefore, testbed system is a popular trend which many research institutes in the world use widely. In Vietnam at present, researching on the performance of wireless network is a current issue and there have been many researches about this. However, most researches are only testing and evaluating the result by using methods proved by mathematics model, or based on the software simulation. Building a testbed system for wireless network is meaningful for basic researches at Vietnam Institute of Information Technology, as well as for extending more co-operations with other outside organizations.

Keywords: Testbed · Experiment · Wireless performance

1 Introduction

Currently, Internet and advanced network technologies, especially wireless network have improved our life in different ways, but the very fast development of network technology also has some limitations. The fact is that researchers and producers of the industry have difficulties in designing and developing the new network technologies. These technologies need to be evaluated and examined thoroughly before being officially applied in finished products.

Simulations allow starting rough evaluation of behavior, productivity and extending ability of a new technology. Therefore, simulation is very useful beginning step (and usually not costly) for testing new technology. However, with the complications in design and measurement, simulations such as NS3 [1] and

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_50

OMNet++ [2], are often used in simplified models, this can easily lead to mistakes of treatment in factual environment. Emulator is an alternative, this combines using the installed prototype and the replaced entity to recreate the behavior in the real world. However, such replacement usually cannot catch up with the complications of factual environment. The result is that only experiment using the installed prototype in reality and in the controlled environment (for example real users, the wireless telecom equipment, the movement of equipment, the layout of bulding, ...) can provide results close to the reality. The backgrounds or means of experiment like this are called testbed.

Thus, network testbed is the major component in developing new network technology, this deep view has resulted in many global scale innovations aiming at designing, providing and managing testbeds. At present, a lot of universities, research institutes in the world have built and been building the laboratory system using their own testbeds, for instance PlanetLab [3] or ORBIT [4]. Testbed is considered an effective replacement for simulations and emulators, where the new technology is evaluated in control but the environment and scale is similar to real life.

In recognition of the important role of testbed in assessing network system, big sponsor of organizations by government for researches have invested considerable resources not only in conducting experiment of testbed hardware, but more importantly in developing new system to support and maximize its usage in global scale. *GENI* (Global Environment for Network Innovations) project [5] is funded by NSF (U.S. National Science Foundation), and *OneLab Future Internet Testbed* [6] project is funded by *FP6 and FP7* (6^{th} and 7^{th} European Union Framework Programs) are typical examples of the recent innovations about this trend.

In comparison with other fields of scientific research, life science for example, a tight cross checking on the experiment results is still absent in networking field. This is due to the fact that even with experiment infrastructure available, there is no clear way to describe an experiment allowing other experiments repeating this. Therefore, in order to increase the rigidness of science in networking field, there should be effective tools and methods to support researching the whole process life cycle. Especially, we need systematic descriptions of experiments, including the resources used and the measurement taken. And the result that is, one person can easily repeat the experiment in similar or different context, and more important capability is to allow others to do the same.

2 Related Works

Emulab [7] is a network testbed, giving researchers a wide range of environments in which to develop, debug, and evaluate their systems. The name Emulab refers both to a facility and to a software system. Emulab is a public facility, available without charge to most researchers worldwide. Emulab provides integrated access to a wide range of experimental environments: An ***emulated experiment*** allows you to specify an arbitrary network topology, giving you a *controllable*,

predictable, and *repeatable environment*, including PC nodes on which you have full "root" access, running an operating system of your choice; ***Live-Internet Experimentation*** which provides users with a full-featured environment for deploying, running, and controlling their application at hundreds of sites around the world; ***802.11a/b/g Wireless testbed*** is deployed on multiple floors of an office building. Nodes are under your full control and may act as access points, clients, or in ad-hoc mode. All nodes have two wireless interfaces, plus a wired control network; and ***Software-Defined Radio*** with USRP devices from the GNU Radio project give you control over Layer 1 of a wireless network – everything from signal processing up is done in software.

PlanetLab [3] is a global research network that supports the development of new network services. Since the beginning of 2003, more than 1,000 researchers at top academic institutions and industrial research labs have used PlanetLab to develop new technologies for distributed storage, network mapping, peer-to-peer systems, distributed hash tables, and query processing. PlanetLab currently consists of 1353 nodes at 717 sites.

NITlab [8] focuses on the design, study and implementation of wireless and wired schemes and their performance in the real environment. In this context, NITlab has developed a facility named NITOS, which stands for *Network Implementation Testbed* using *Open Source* platforms. NITOS facility currently consists of over 100 operational wireless nodes and is designed to achieve reproducibility of experimentation, while also supporting evaluation of protocols and applications in real world settings. NITOS facility is geographically separeted in 3 deployments. The Outdoor one at the exterior of the University of Thessaly (UTH) campus building, the Indoor one at the basement of the UTH's building and the Office testbed deployed at CERTH's office building in Volos.

Most of testbed described above use OMF (cOntrol and Management Framework) [9] a control, measurement, and management framework for testbeds. OMF was originally developed for the ORBIT wireless testbed at Winlab, Rutgers University [10]. Through active development and extensions at NICTA, it has now evolved into an open source framework, which supports heterogenous wired and wireless resources. OMF is one of the few candidates currently being evaluated as potential testbed framework by both the GENI and the Onelab initiatives. It is currently deployed and used on different testbed in Australia, USA [4], and Europe [8], with many of them in active use 24/7.

3 Building Testbed System Based on Customized Wireless Nodes

The objective of our testbed system includes: building wireless testbed system for wireless network research at Institute; Evaluating some solutions for improving network performance based on testbed, and provides testing and experimenting services for external users. Figures 1 and 2 illustrate the corresponding overview diagram and logic diagram of testbed system. With this system, users use *experiment description language* (Ruby language source code) to describe the

experiment components, the resources needed to run experiments, running time, related parameters... After that, description script will be sent to *Experiment Controller* (EC) to execute script, EC communicates with *Resource Controller* (RC) to request resources for running experiment that described in script. RC will allocate resource to run experiment, return results to EC, and users can analyze, evaluate, and process experiment outputs. With each user A, B, ... the operations of describe experiment, request resources, run and get experiment results will be repeated.

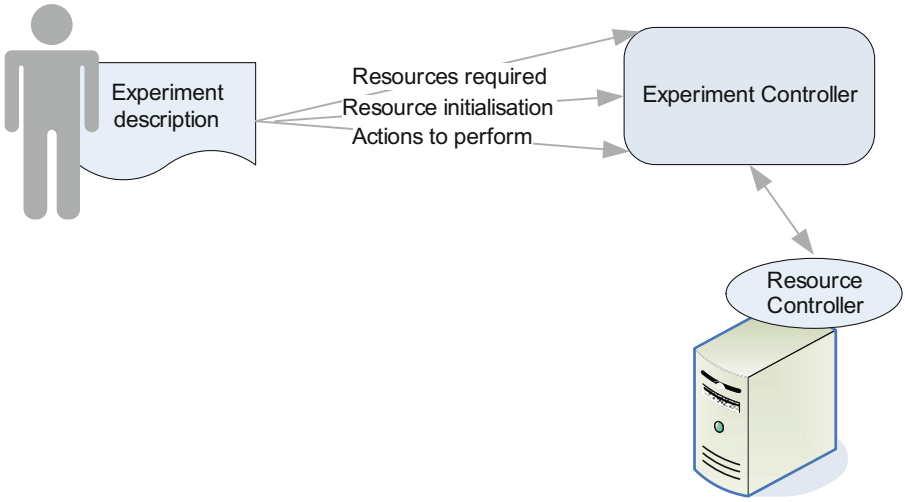


Fig. 1. Overview diagram of IOITtestbed.

Figure 3 is installation diagram of IOITTestbed. To manage control information, for example, transferring the information of description script from wireless nodes to Resource Controller and Experiment Controller servers, we use wired connection to increase processing speed. The connection between wireless nodes includes two interfaces: wired interface use for connecting to servers, and wireless interface use for connection nodes-to-nodes to run experiments. And we will have extension modules for wireless sensor, OpenFlow, ZigBee networks... in the future. This expansion can be done easily because the customized wireless nodes was ordered by our own demand.

4 Evaluating Network Performance Based on Wireless Testbed System

4.1 Theory of Maximum Throughput

In wireless networks, the maximum throughput is always much smaller than channel data rate, that caused by the overhead by header information (see Fig. 4)

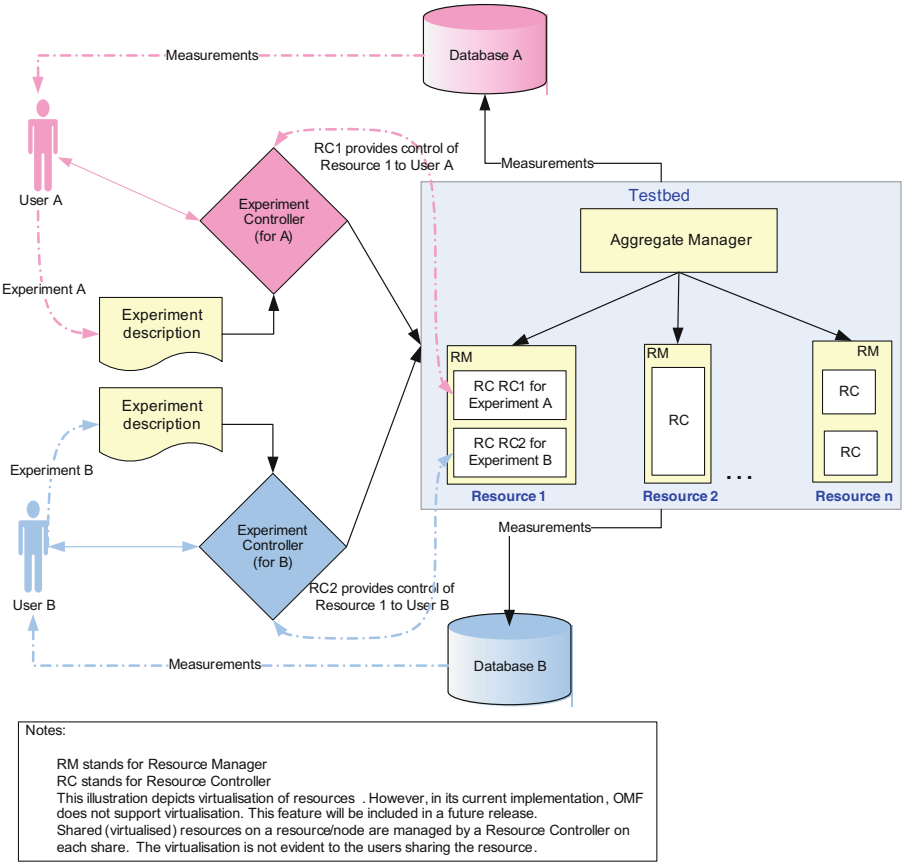


Fig. 2. Logic diagram of IOITtestbed.

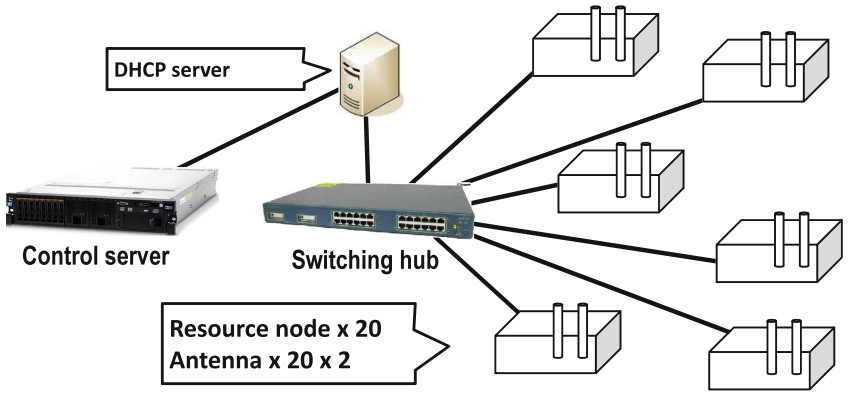


Fig. 3. Installation diagram of IOITtestbed.

and by the operation of IEEE 802.11. To calculate maximum throughput, we need carefully analyze the overheads associated with access mechanism as below.

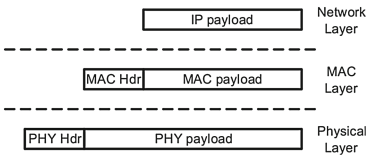


Fig. 4. Encapsualation overheads.

The basic access mechanism is a two-way handshaking technique. This mechanism is characterized by the immediate transmission of an ACK frame by the destination station, upon successful reception of a packet transmitted by the sender station.

Equation 1 defines the maximum expected throughput for a single active session (i.e., only a sender-receiver couple active) when the basic access scheme is used:

$$Th_{Basic} = \frac{P}{DIFS + \delta + T_{(H+P)} + SIFS + \delta + ACK + \sigma CW_{min}/2} \quad (1)$$

where

- P is the packet payload size,
- δ is propagation delay,
- T_{H+P} is the time required to transmit a data frame, this includes the PHY header, MAC header,
- ACK is the time required to transmit an ACK frame, this includes the PHY header
- $\sigma CW_{min}/2$ is the average back off time.

4.2 Simulation and Evaluation

We make a simulation with single-hop as in Fig. 5 to evaluate our analysis in maximum throughput with Eq. 1. Where source station $M1$ sends data to destination station GW. The values of the parameters used to obtain numerical results, for both the analytical model and the simulation runs, are summarized in Table 1.

The simulation results by NS-2 [11] are in Tables 2, 3.



Fig. 5. A basic single-hop wireless network model.

Table 1. Simulation parameters

Parameters	Values
MAC header	34 Bytes
PHY header	24 Bytes
ACK	14 Bytes
Propagation delay	1 μs
Slot time	50 μs
SIFS	28 μs
DIFS	128 μs

Table 2. The analysis and simulation results in maximum throughput (Mbps) of 802.11b.

Parameters		Throughput	
Channel rate (Mbps)	Payload (B)	Analysis	Simulation
2	1024	1.64	1.42
2	512	1.40	1.16
11	1024	5.03	3.53
11	512	3.26	2.26

Table 3. The analysis and simulation results in maximum throughput (Mbps) of 802.11g.

Parameters		Throughput	
Channel rate (Mbps)	Payload (B)	Analysis	Simulation
11	1024	8.78	8.25
11	512	7.31	6.90
54	1024	24.10	23.34
54	512	15.51	15.50



Fig. 6. Two wireless nodes in ad hoc mode.

Table 4. TCP/UDP parameters for iPerf.

Parameter	Value
TCP window size	85 KB
TCP data rate	1.54 Mbps
UDP buffer size	208 KB
UDP data rate	2.18 Mbps

4.3 Evaluation with IOITtestbed

We define the similar experiment with testbed as follow: two wireless nodes in Fig. 6 are configured in ad hoc mode, we use iPerf [12] (was compiled follows specifications of testbed application) to generate data for sender and measure data received for receiver.

The commands for sending and receiving data with iPerf is defined as:

- Receiver (server): */usr/bin/iperf-oml2 -s -i 1 -u*
- Sender (client): */usr/bin/iperf-oml2 -c 172.16.3.7 -t 300 -u -b 200M*

The values of TCP/UDP parameters are summarized in Table 4.

The results with 300 s running time: there are 444229 sent and received packets (datagram). The average throughput is 12.3 (Mbps) with TCP and 17.4 (Mbps) with UDP. Figure 7 illustrates detailed comparision of TCP and UDP throughput in sending and receiving processes.

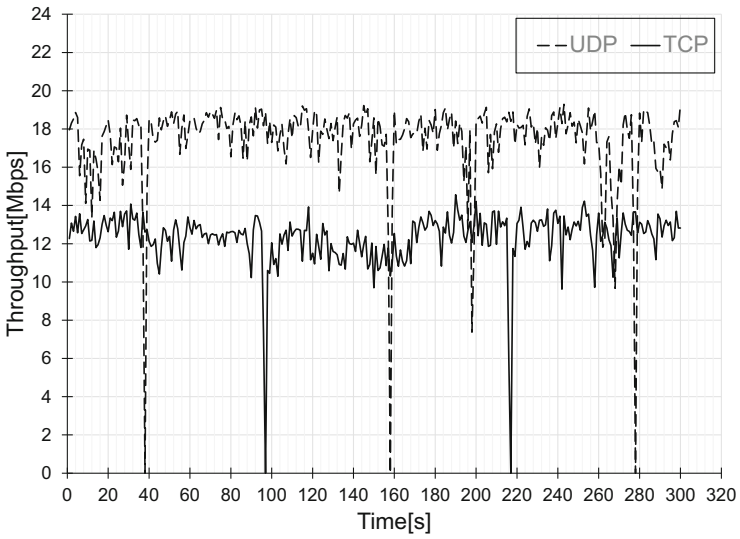


Fig. 7. TCP and UDP throughput.

Thus the results of testbed-based evaluation relatively close to theoretical results. When the UDP throughput is always higher than TCP throughput, because TCP uses window size and other flow control, congestion control mechanism, so throughput will be decreased, while UDP uses "best-effort" mechanism which ignored control information, for the purpose of sending data as much as possible.

5 Conclusion

Wireless network is increasingly becoming important infrastructure for family, business organizations, even to the extent of the industry. However, the technology used in wireless network has to be tested, examined and evaluated before being released and applied officially. In the past, researches on wireless network technology were mainly examined and evaluated on the basis of mathematics model, or the simulation tools. The advantage of these solutions is the equipment is not costly because of proving with mathematics or using software to write testing script, analyze results, ... The disadvantage is that this is limited by the ideal conditions and theories can be assessed because modeling or simulating cannot reflect the real factors of network environment. The trend of using testbed to evaluate network parameters is more advantageous than the modeling and simulating methods. In today condition of Vietnam, there has been very few or even no research team using testbed. The project of building testbed for wireless network at our Institute of Information Technology has focused on building a networking laboratory with new approach. At first stage, we have successfully carried out testbed building on the basis of facilities and equipment invested. Based on that, we have some basic parameters for the network. In the following stages, we will complete testbed with capability of providing services for the users, and try to test some proposed problems of improving network performance.

Acknowledgement. This work was financially supported by National Key Lab for Networking Technology and Multimedia (2015–2016) with the project "Building a wireless testbed system for evaluating network performance.".

References

1. NS3 Network Simulator. http://www.nsnam.org/
2. OMNet++ Simulator. http://www.omnetpp.org/
3. PlanetLab Consortium, Planetlab: An open platform for developing, deploying, and accessing planetary-scale service. http://www.planet-lab.org/
4. Raychaudhuri, D., Seskar, I., Ott, M., Ganu, S., Ramachandran, K., Kremo, H., Siracusa, R., Liu, H., Singh, M.: Overview of the ORBIT radio grid testbed for evaluation of next-generation wireless network protocols. In: Wireless Communications and Networking Conference, vol. 3, pp. 1664–1669. IEEE, 13–17 March 2005

5. Clark, D., Shenker, S., Falk, A.: GENI: Global Environment for Network Innovations. http://groups.geni.net/geni/rawattachment/wiki/OldGPGDesignDocuments/GDD-06-28.pdf
6. OneLab: Future Internet Testbed. https://onelab.eu/
7. White, B., Lepreau, J., Stoller, L., Ricci, R., Guruprasad, S., ewbold, M., Hibler, M., Barb, C., Joglekar, A.: An integrated experimental environment for distributed systems and networks. In: ACM SIGOPS Operating Systems Review - OSDI 2002: Proceedings of the 5th Symposium on Operating Systems Design and Implementation, vol. 36, pp. 255–270, Winter 2002
8. The Network Implementation Testbed Laboratory. http://nitlab.inf.uth.gr/NIT/NITLab
9. Rakotoarivelo, T., Ott, M., Jourjon, G., Seskar, I.: OMF: a control and management framework for networking testbeds. ACM SIGOPS Oper. Syst. Rev. **4**, 54–59 (2011)
10. Ott, M., Seskar, I., Siraccusa, R., Singh, M.: ORBIT testbed software architecture: supporting experiments as a service. In: First International Conference on Testbeds and Research Infrastructures for the Development of Networks and Communities, Tridentcom 2005, pp. 136–145. IEEE, 23–25 February 2005
11. NS2 Network Simulator. http://www.isi.edu/nsnam/ns/
12. iPerf - The network bandwidth measurement tool. https://iperf.fr/

Predicting Early Crop Production by Analysing Prior Environment Factors

Tousif Osman, Shahreen Shahjahan Psyche, MD Rafik Kamal, Fouzia Tamanna, Farzana Haque, and Rashedur M. Rahman(✉)

Department of Electrical and Computer Engineering, North South University, Plot-15, Block-B, Bashundhara, Dhaka, Bangladesh
{tousif.osman,shahreen.psyche,rafik.kamal,fouzia.tamanna, rashedur.rahman}@northsouth.edu, soniya.farzana@gmail.com

Abstract. Bangladesh has an agriculture dependent economy and hence prediction of agricultural production is of great importance to us. In this research we develop a model that considers and analyzes weather and climate prior to specific crop plantation and maps a correlation between these two. It allows us to provide information about the crop state, in quantity and quality with the possibility of early warnings so that timely interventions can be undertaken. The approach advocated in this paper is to help the people with food security and early warning system.

Keywords: Data mining · Adaptive learning · Machine learning · Prediction · Agriculture · Soft computing · Environment

1 Introduction

The economy of Bangladesh is highly depended on the agriculture but agriculture faces a lot of challenges each year. One of the problems agriculture face is crop yield prediction. Primarily the crop yield depends on weather, planning of harvest operation, diseases and pests. It is important for the farmer to decide which crops will have the best production before the harvest takes place. Timely and accurate crop yield prediction helps the farmer to ensure maximum production of crops for present and future. We have constructed a model in this research considering environmental variables on which crop production depends directly and indirectly. It uses the value of sunshine, cloud, wind-speed, humidity, temperature and rainfall for computing the production of certain crop. Extracting knowledge from this raw data is very difficult. But one can get the knowledge of the data to predict the major agricultural production which is depended on climate and weather by data mining methods and techniques. So, we have designed our system by which farmers can produce more crops.

2 Related Work

A good number of researches has been done to predict crop yield. We have studied and incorporated ideas and methods from those researches in to our

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_51

system. In this section we will mention few research works and their approaches those we have considered while doing our research and building the system.

In paper [1], authors have found the yield of crops of Bangladesh using regression tree, neural network, ensemble learning, and linear/nonlinear regression. To find the similar characteristic of environment between the regions they have used k-means clustering and self-organizing map (SOM) and they mentioned two prominent regions. After doing this, they have taken average of 4 month's weather attributes values as input and yield of rice as output of the same time. In next research [2], the authors described Naive Bayes techniques for classification of agricultural land soils. The research utilized data collected from seven commonly occurring soil types and analyzed with agricultural data set using data mining techniques. The Naive Bayes helped to classify the soil based on texture of soil profiles and also classification worked with a simple probabilistic classifier based on independence assumption. In paper [3] the authors have mainly pointed on data mining techniques to predict major crop yield with existing data. The authors in [4] proposed a system named GZ-Agri GS. The local farmers can easily gain the knowledge about scientific guidance for crop management decisions from the online service. This research uses Domain knowledge and model with GIS technology to gain useful results. In the research paper [5], the authors discussed a system that is applicable for the real-life data quality control and assurance of reliable and error-free agro-meteorological data. The authors design a prototype system for detecting abnormal weather prediction.

3 Theory

The main goal of this research is to predict the production of some major crops of Bangladesh for upcoming season. We have used 10 years monthly environmental data and crop production data to model our system. Major components of the research have been described in the following three subsections.

3.1 Prediction Hypothesis

We want to construct a system that depend on prior environmental parameters (i.e., sunshine, rainfall, soil salinity). Yield of crops depends on present environment parameters. However if we consider the fact, most of the environmental parameters are depended on each other and previous environmental parameters we can directly correlate crop yield with prior seasons environment. This is our research's basis on which we have constructed our system. We will prove the hypothesis in our research.

3.2 Data Attributes

Six components of regional weather data of Bangladesh is archived and publicly available. They are temperature, rainfall, humidity, sunshine, cloud coverage and wind speed as monthly average of last 60 years. We have collected this data from

Bangladesh Agricultural Research Council (BARC) [6]. Data for 6 crops were publicly archived by Bangladesh bureau of Statistics (BBS) [7]. Those are Aus Rice, Amon Rice, Boro Rice, Jute, Potato, and Wheat, having 2313 entries of last 10 to 8 years. All the data attributes are described briefly below:

Temparature is an attribute of weather. It is a vital factor for the production of crops. Its meta attributes are shown in Fig. 1. During the model generation we have considered past 3 months weather data from the plantation month of specific crops. So in the data preprocessing stage 3 relevant consecutive months' attributes has been selected. **Rainfall** is another vital attribute for production of crops. Rainfall and all other environmental data has same table structure as temperature but the rainfall measurements are taken in millimeter. **Humidity** is the next attribute we have considered. This is the parameter of water vapor in the atmosphere. This is also an important factor for production. Here all the measurement is done in percentile. **Sunshine** is an attribute of weather. Sunshine plays a significant role for the production of crops. All the measurement of here has been done as sunshine hours. **Cloud Coverage** is one of the environmental parameters. It does not directly effect crop yield but it has indirect impact according to our in Hypothesis Sect. 3.1. The measurement has been done in Oktas. **Wind Speed** is one other parameter of weather. Like rest of the environmental data it has a similar table structure. All the measurement of temperature has been done in Knot.



Fig. 1. Temperature meta data on the database.

Crop Area is one of the crop attributes. Crop data for all 6 crops has similar table structure. Crop table has several attributes. Among them there are few that represent the area of crop cultivation lands but in different units (i.e. Acres, Hectors etc.). We have considered only one attribute, Area Hectors (Ahectors). One of the important issues while predicting the production is the area of the land which is used for production. This represents the total area under a certain region. Figure 2 is the data table for the crop Jute. Table structure for the rest of the crops is similar. **Crop Production** is our primary attribute. Our research goal is to predict crop production for unseen data. This represents the production of previous years on different regions. In Fig. 2 we can see that there are few attributes that represent production but in different units (i.e. Ton, Maunds), like crop area. Among them we have selected production in Metric Ton (M.Ton). **Region** is basically one of the crucial attributes. Bangladesh has been divided into separate regions. All the calculation will be done based on this region as yield varies based on region. We have analyzed and produced model and data representation based on separate regions. **Production Year** is the year of production for a crop. Here this attribute represents different years. For the research we have taken production year from 2007 to year 2013.



Fig. 2. Jute data table in the database.

3.3 Training the Model

One of the main objectives of this research is to construct a learning model for crop yield. We have constructed the system to predict production in terms of numerical data (M.Ton). We have applied Linear Regression and Neural Networks to train our the model. Finally we have compared methods in terms of error rate and prediction accuracy. As both of those models operate on numerical data only we cannot consider regions, which is a categorical data. But we cannot ignore the region while predicting crop yield. To resolve this we have designed the system to generate smaller models for all regions.

Linear Regression(LR) is a mathematical term that is used in statistical measurement to determine the relationship between a dependent variable and to its corresponding one or more independent variables. In data-mining linear regression is a numerical procedure to forecast the outcome of some dependent variable. **Neural Network**(NN) in Computer Science, is a computational model that utilizes the notion of back propagation to design a model. The number of hidden layers by default depends the number of attributes and also the number of classes for the given data set. If number of hidden layer is μ, number of attributes is α and number of classes is β then, the number of hidden layer in a neural network by will be:

$$\mu = \left[\frac{\alpha + \beta}{2}\right] + 1 \tag{1}$$

Machine Learning algorithm has been used improve the performance of the models. We have used Bootstrap Aggregating (Bagging) in our research. It is a machine learning algorithm which learns by splitting the training set. Bagging trains n number of models on the basis of new n numbers of training set. The algorithm will take the average and produce the final result. In our system, the value of n is 5 and data has split ratio 9:1.

4 Methodologies

We have used the concepts explained in the theory section to construct our system. Initially data is loaded from the database and preprocessed all the separate data tables to create a single table. Afterwards data is passed through a loop and splitted with respect to regions. This smaller dataset then is passed through Learning and bagging algorithms to train the model. Performance testing and prediction of unseen data is done in the loop. Finally all the resultant data are put together at the end of the loop and returned as the final result. The system has been built using RapidMiner Studio. Figure 3 shows major components of the main process of our system. All the major components of the system and its working will be explained in this section.

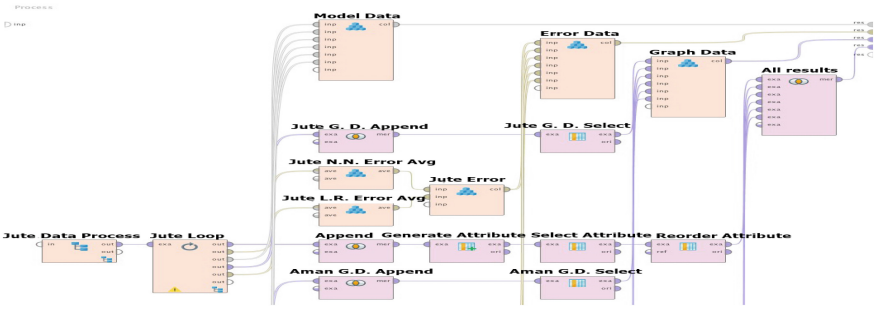


Fig. 3. Partial Design of the System.

4.1 Data Preprocessing

First step of the system is the data preprocessing. In the Fig. 3 the left most block named as Jute Data process is the starting point of the system. This block is a sub process which preprocesses the data. Figure 4 shows major components of this process. This block preprocesses jute data only. At the beginning all separate data tables are loaded and select operators are used to select the proper weather attributes. In case of Jute, it is planted on the month of March so the region, year and January to March weather data attributes are selected. Then those weather attributes are renamed and given a generic name and all weather tables are joined together. Finally the crop production table is joined with the previously joined table and Set Role operator has been used to identify which data attribute we want to predict.

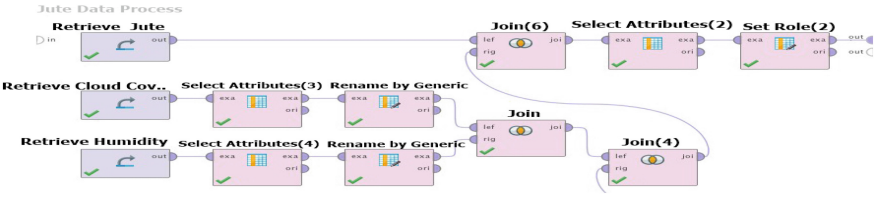


Fig. 4. Components of Data preprocessing sub process.

Figure 5 shows the final joined table for jute after preprocessing the data. In this sub process loading and joining of all the weather data has not been shown as it is a repeated work. Rest of the crops has similar preprocessing blocks that are not shown in Fig. 3. Those blocks also work in the same way but they operate on relevant data tables and attributes.

M.Ton	Year	Region	AreaH	cloud_m1	cloud_m2	..	..	..	humid_m3

Fig. 5. Preprocessed Jute data table.

4.2 Looped Subprocess

After preprocessing, the data is fed in a looped sub process. As explained in the Training Model Sect. 3.3. In the Fig. 3 we can see this loop block for jute named Jute Loop just after the Jute Data sub process. This loop iterated over all unique values of the Region values that is with all the region names. Figure 6 shows the loop process flow for jute preprocessed data table.

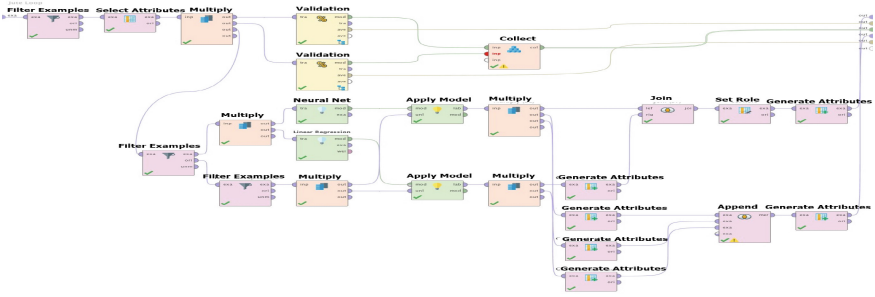


Fig. 6. Loop process for Jute.

For each iteration of the loop the name of a region is made available as a macro. Inside Loop process data is cloned in two data sets to perform testing and performance generation and Predict an unseen data in this case 2009 production of each crops. At first preprocessed data is split regionally and categorical attribute will be removed from the sub-tables because one table corresponds to only one region and as the methods we are using to train the model accepts numerical data only. After that sub-tables are cloned into two sets. One set is used to train two separate model using LR and NN. After training the model crop yield is predicted. Figure 7 shows the NN for Jute. NN model has one hidden layer. Both of the model are wrapped inside Bagging block which applies Bagging Algorithm. Another one of the cloned data sets are feed in the testing branch of Fig. 6. Tenfold method is used to measure the performance of the model. Finally the performance result of both algorithms are grouped in and collected in a data collection. At the end of the loop the categorical regional attribute is regenerated and added to the resulting table.

4.3 Data Collection and Produce Result

From Fig. 3 we can see several data collector is used to collect the results. During this step resulting data is processed so that we can plot and compare the results.



Fig. 7. Neural Network model for "Jute".

5 Results and Data Representations

In this research we have worked with six crops and applied our model on all six crops using both methods. In this section results and findings of our research has been described.

5.1 Testing Results

We have applied Root Mean Squared Error (RMSE) method using split testing algorithm for error calculation. Error rate of all crops using both methods has been listed and compared in the following table (Table 1).

5.2 Tabular Prediction of Result by Production on 2009

To test the result of our system we removed the data of 2009 while training the model and the predicted the result. Figure 8 shows the output of few crops as tabular format for both prediction models.

Table 1. Result Error

Crop	*Neural Net RMSE*	*Linear Regression RMSE*
Jute	1.943 +/- 0.00	0.992 +/- 0.00
Aman	0.248 +/- 0.078	0.334 +/- 0.202
Aus	0.106 +/- 0.00	0.173 +/- 0.00
Boro	0.236 +/- 0.00	0.244 +/- 0.00
Potato	8.136 +/- 0.00	2.983 +/- 0.00
Wheat	0.363 +/- 0.00	0.318 +/- 0.00

M.Ton	Neural Net	Linear Regr...	Crop Name	Region	Year
3.309	3.495	3.479	Boro Rice	Bhola	2009
4.009	4.153	3.991	Boro Rice	Jessore	2009
2.326	2.016	2.026	Boro Rice	Patuakhali	2009
3.987	3.986	3.922	Boro Rice	Bogra	2009
3.777	3.895	3.893	Boro Rice	Dinajpur	2009
13.172	10.840	12.330	Potato	Cox's Bazar	2009
20.392	16.583	14.369	Potato	Comilla	2009
19.245	19.653	16.163	Potato	Dhaka	2009
17.942	17.687	17.347	Potato	Barisal	2009
11.120	18.618	19.831	Potato	Bhola	2009
0	13.687	13.450	Potato	Jessore	2009
0	19.854	31.455	Potato	Patuakhali	2009
11.673	15.194	14.822	Potato	Bogra	2009
11.183	12.847	15.649	Potato	Dinajpur	2009

Fig. 8. Jute prediction result for year 2009.

5.3 Bar Charts of Results by Production on 2009

We have represented all results of all region as bar chart for year 2009 in Figs. 9, 10, 11, 12, 13 and 14.

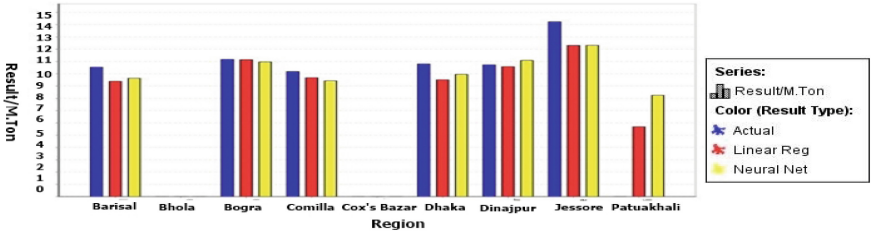


Fig. 9. Jute Prediction for 2009.

5.4 Research Finding

From all three research results, RMSE, Tabular result and Graphical we can see that crop production can be predicted considering prior environmental attributes with good accuracy. RMSE of all crops is relatively low and hence this satisfies our hypothesis. Furthermore from the comparison we can see that NN provides slightly better results. However NN takes more time to train the model.

6 Future Work

In this research we have considered 6 attributes to predict the yield of a crop. However correlation of environment attributes has little dependency with the

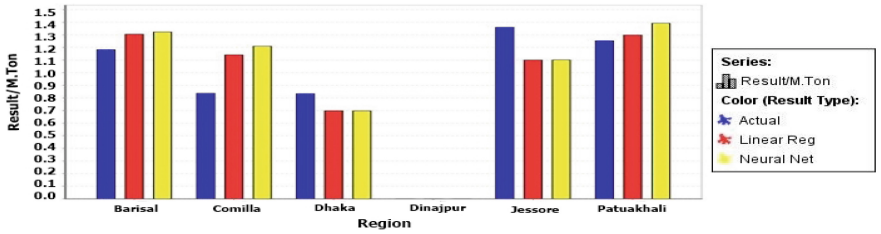


Fig. 10. Aus Prediction for 2009.

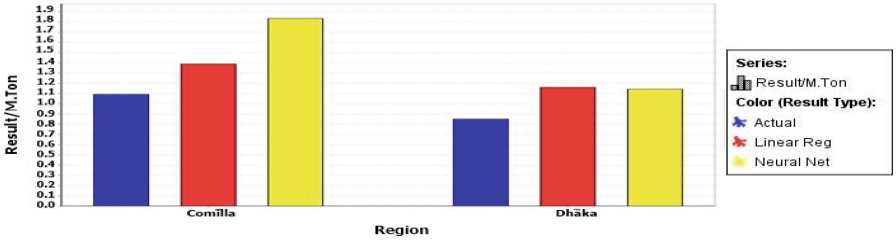


Fig. 11. Amon Prediction for year 2009.

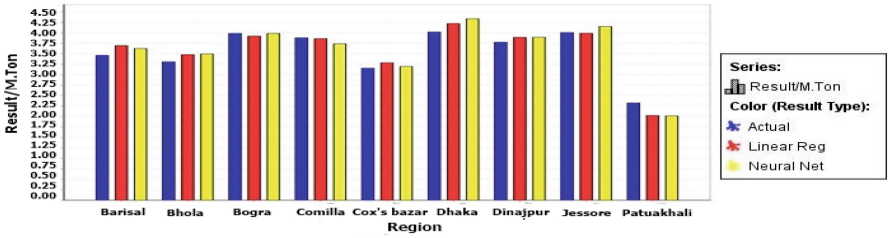


Fig. 12. Boro Prediction for year 2009.

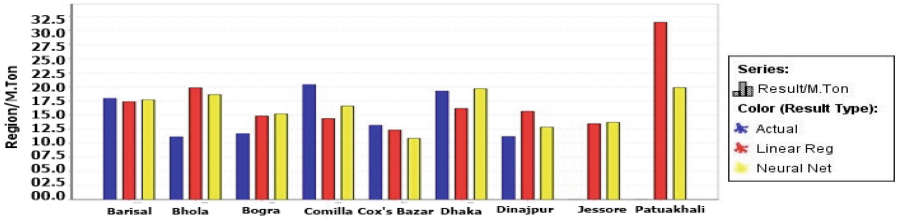


Fig. 13. Potato Prediction for year 2009.

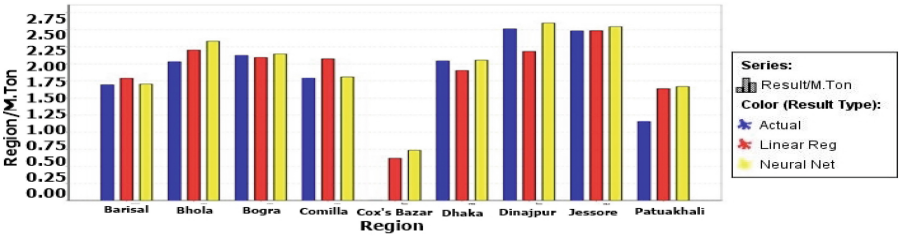


Fig. 14. Wheat prediction for year 2009.

attributes of previous months. We plan to collect more environment attributes and integrate in the system. Currently we are not considering already identified factors for predicting crop production that are used by farmers and environmentalist. We plan to include that knowledge in the system for better prediction.

Acknowledgments. We would not complete our research paper without having the support of the organizations named Bangladesh Agricultural Research Council and Bangladesh Bureau of Statistics. We would like to extend our sincere gratitude to those organizations.

References

1. Rahman, M., Haque, N., Rahman, R.: Application of Data Mining Tools for Rice Yield Prediction on Clustered Regions of Bangladesh. In: 17th IEEE International Conference on Computer and Information Technology (ICCIT), pp. 8–13. Daffodil International University, Dhaka, Bangladesh (2016)
2. Bhargavi, P., Jyothi, S.: Applying Naive Bayes Data Mining technique for classification of agricultural land soils. In: International Journal of Computer Science and Network Security, pp. 117–122 (2009)
3. Ramesh, D., Vardhan, B.: Data mining techniques and applications to agricultural yield data. Int. J. Adv. Res. Comput. Commun. Eng. **2**, 3477–3480 (2013)
4. Sha, Z., Zhang, M.: Development of web-based decision support system for field-based crop management. In: Geographic Information Systems, pp. 1–4 (2007)
5. Mateo, M., Leung, C.: Design and development of a prototype system for detecting abnormal weather observations. In: Proceedings of the 2008 C3S2E conference on - C3S2E 2008 (2008)
6. Bangladesh Agricultural Research Council (BARC)-Government of the People's Republic of Bangladesh, Barc.gov.bd (2016). http://barc.gov.bd/. Accessed 01 May 2016
7. Bangladesh bureau of Statistics (BBS). http://www.bbs.gov.bd/. Accessed 01 May 2016

Prediction of Generalized Anxiety Disorder Using Particle Swarm Optimization

Wahidah Husain[1(✉)], Saw Hui Yng[1], Nur'Aini Abdul Rashid[2], and Neesha Jothi[1]

[1] School of Computer Sciences, Universiti Sains Malaysia, Penang, Malaysia
wahidah@usm.my, {shyng.ucom12,nj14_com042}@student.usm.my
[2] College of Computer and Information System, Department of Computer Sciences, Princess Noura bint Abdulrahman University, Riyadh, Kingdom of Saudi Arabia
nurainipng@gmail.com

Abstract. Diseases can be predicted by using historical patient information stored in clinical databases. Large data is required to ensure the accuracy of prediction. However, processing and extracting valuable information from huge data is a challenging and time-consuming task. Missing and incomplete data may easily cause the data to be ignored and not fully utilized in the prediction. In this paper we focus our study on generalized anxiety disorder. Prediction of generalized anxiety disorders is carried out using feature selection and classification approach. This research focuses on studying and implementing Particle Swarm Optimization algorithm and Fuzzy Rough Set in the classification of generalized anxiety disorders. Performance of classifier model is evaluated respectively based on the accuracy, sensitivity and specificity of results produced. It is found that the proposed hybrid approach in feature selection has different results in performance depending on the selection of classification technique.

Keywords: Data mining · Fuzzy rough set feature selection · Generalized anxiety disorder · Particle swarm optimization

1 Introduction

Everyone commonly experiences anxiety, which it is one of the normal human emotions. Anxiety is the body's natural response to danger, an automatic alarm that goes off when you feel threatened, under pressure, or are facing a stressful situation. However when anxiety is constant or overwhelming, it will results in anxiety disorders which cause distress that interfere with a person's ability to lead a normal life. With gradually increasing number of anxiety disorders around the world and increasing cost of healthcare organizations, screening of anxiety disorders is required in health care field in order to save time, money and human lives. By having more efficient and accurate screening mechanism, it will guide physicians in screening among the patients for anxiety disorders faster and more accurate.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_52

In recent years, medical data is growing exponentially and often valuable information is hidden in the data. Massive amount of data is making it difficult for human to process and analyze the data as it may take weeks to obtain valuable information from raw data. Hence, data mining becomes an important tool in processing big data. In this study, data mining methods are being applied in the screening of anxiety disorders. Data mining is used to discover knowledge from historical data of various domains. In healthcare domain, data mining can be used to mine medical data as it is able to process huge amount of data about patients, diseases, diagnosis and medicine.

Feature selection is the process of choosing a subset of relevant and optimal features for model building. Raw data may contain redundant and irrelevant features. Without removing those features, it would affect the performance of the classifier model. However, it is usually a difficult task in determine which features are useful as there would exist a complex interaction among features. Besides, the problem of large search space makes feature selection process more difficult to implement. Therefore, by suggesting global search technique based on swarm intelligence approach, it could overcome the feature selection problems and may results in better classification performance.

There is not much research that has been done on anxiety disorders using swarm intelligence approach. Swarm intelligence (SI) is an innovative intelligent optimization-based technique which is inspired by the biological behavior of birds [1,6]. It has primarily been used and found very efficient in traditional optimization problems [3]. However in healthcare field especially for anxiety disorders, there is still lack of research of implementing swarm intelligence on it. Among the swarm intelligence approaches, Particle Swarm Optimization (PSO) algorithm is implemented in this study to optimize the feature selection in order to classify people with and without anxiety disorder effectively.

2 Background Study and Related Work

Tremendous amount of health care data has flooded from every health organization. Explosive growth in stored health care data urged the need for a tool that able intelligently transforming vast amount of data into useful information and knowledge. Data mining is used to find patterns and regularities in large number of datasets that are difficult to analyze manually. It helps to predict future trends and behaviors that allow people to make knowledge-driven decisions.

2.1 Feature Selection

Feature selection is a process of selecting subsets of features, which are relevant for classification. It is also called variable selection or attributes selection. It reduces dimensionality of data by selecting only a subset of relevant features to model construction, which optimize the accuracy of the classifiers. Data may contain relevant, irrelevant, redundant and noisy features.

A wrapper feature selection algorithm using PSO and Support Vector Machine (SVM) is proposed by Lin et al. [4]. The proposed algorithm is found that it could optimize the parameters in SVM and search for the best feature subset simultaneously. Meanwhile, Chakraborty [7] carried out a comparative study on the performance of PSO and Genetic algorithm (GA) in a filter feature selection algorithm with a fuzzy set based fitness function. The results show that PSO performs better than GA in classification. Ming [8] has proposed a feature selection method based on ACO and rough set theory. Forward selection is applied into the proposed method to search for best feature subset. The results showed that the proposed method performs better in classification with fewer features. Gao et al. [9] proposed a wrapper feature selection algorithm based on ACO, which used to study network intrusion detection. Based on the related works, there are numerous methods of feature selection available nowadays. However in this study, we are focus on Fuzzy Rough set feature selection method and it is discussed in the following section.

2.2 Fuzzy Rough Set Feature Selection

Fuzzy Rough set is a hybrid feature selection method that applies rough set theory and fuzzy set theory. Rough set theory [10] and fuzzy set theory are generalized of classical set theory that deals with vagueness and uncertainty in data analysis. In rough set theory, the attribute reduction algorithm eliminates redundant features and chooses a subset of features that has same discernibility with original set of features [10]. In data analysis, rough set theory has providing efficient algorithm in discovering hidden pattern in data and also generating sets of decision rules from data that is concise and valuable [10]. Although rough set theory is an efficient tool for managing problems in feature selection, it is found that it is unable to handle real-valued features. In the data, some values of features may be real-valued and it cannot be handled by traditional rough set. Besides, rough set theory has the limitation of handling only nominal data. It does not able to handle continuous data [10]. Therefore, fuzzy set theory is combined with rough set theory to overcome issues of feature selection using only rough set theory.

2.3 Particle Swarm Optimization

Particle Swarm Optimization (PSO) algorithm was proposed by James Kennedy and Russell C. Eberhart [11] in 1995. It is an evolutionary computational technique based on swarm intelligent. It was inspired by the social behavior of bird flocking and fish schooling. It is an AI technique that can be used to find approximate solutions to extremely difficult or impossible numeric maximization and minimization problems. It is also suitable for search of optimal solutions and based on the concept of swarm. Therefore, it has been used for approaches that can be used across a wide range of applications and fields such as in business, medical and biological fields.

Sousa et al. [12] proposed the first study on the application of PSO to classification. It proposes the use of the particle swarm optimizer as a new tool for data mining. Three variants of PSO algorithms namely discrete PSO (DPSO), constricted PSO (CPSO) and linear decreasing PSO (LDPSO) are applied in the study. Genetic algorithm (GA) and tree induction algorithm (J48) are also applied for performance comparisons with PSO. Classification rule discovery algorithm, sequential covering algorithm and validation algorithm are employed. The results show PSO proved to be a suitable candidate for classification tasks. It is evaluated that PSO can obtain competitive results against GA and J48 algorithms in the data sets used.

Besides, Holden and Freitas [13] proposed a hybrid PSO and ACO algorithms for hierarchical classification. Its goal is to build rule-based classification models that can handle both nominal and continuous variables. The proposed PSO/ACO algorithm is tested on biological data set which consists of hierarchical functional classification of enzymes. The result shows the hybrid PSO/ACO algorithm is having better results with respect to classification accuracy. Besides, it also achieves better results in terms of comprehensibility of the discovered rule set.

3 Methodology

A graphical development methodology is shown in Fig. 1. The research methodology consists of 4 phases: (A) data acquisition, (B) data pre-processing, (C) model building and (D) performance evaluation.

In the data acquisition phase, an online-based survey is conducted. After collecting sufficient amount of data from the survey. The raw data is then organized and pre-processing for data cleaning in the second phase. The preprocessed data then undergoes model building stage in which three approaches are used to build model for every classifier. The first approach is building the model using only basic learning classifiers without implement feature selection and PSO. The classifiers used in this study are Naïve Bayes, Support Vector Machine (SVM), K-Nearest Neighbor (KNN) and decision tree. Next, the second approach is model building using classifiers with PSO, but without going through feature selection. The last method is using classifier with PSO and feature selection. For this approach, feature selection is carried out by using fuzzy rough set feature selection approach. Any unimportant, redundant and noisy features can be removed by feature selection. Therefore, it would reduce the number of iteration that PSO carried out in feature selection to search for optimal feature subset. Finally, the performance of the model is then evaluated in last phase based on confusion matrix. Based on the values in confusion matrix, accuracy, sensitivity and specificity of the model can be evaluated and thus can be compared among other models.

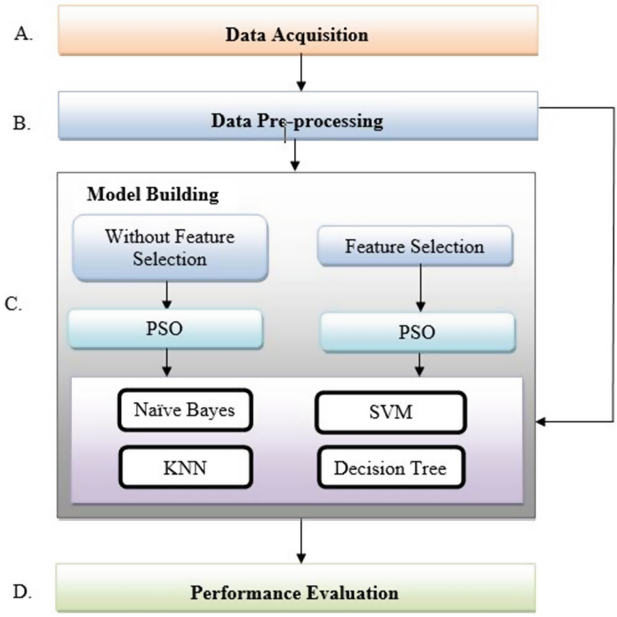


Fig. 1. Research methodology

4 Implementation

Experimental work is carried out by developing a prototype using NetBeans in Java programming language. The API of WEKA Waikato Environment for Knowledge Analysis) software is imported into Netbeans project by adding weka.jar file into the library of Java program. The anxiety disorders dataset consisted of 26 attributes as shown in Tables 1 and 2 with 183 instances. Four classification techniques namely Naïve Bayes, SVM, KNN and decision tree are used to perform classification on the anxiety disorders data. For every classifier model build in this study, ten-fold cross validation method is used to test the performance of the classifiers. In order to study the impact of feature selection and PSO on anxiety disorders, the dataset are processed and mined using three approaches. The first approach is building the model by only using the classifier. Second approach is building the model by applying PSO and classifier. The last approach is model building by implementing feature selection, PSO and classifier on the dataset. For the feature selection, the supervised attribute selection Fuzzy Rough Set method is chosen in this study.

4.1 Attributes Description

The dataset contains 26 attributes. The attributes can be divided into two categories: general attributes (Table 1) and attributes based on Beck Depression

Table 1. General attributes

No.	Attribute	Type
1.	Gender	Nominal: Male or female
2.	Age	Numeric: Age in years
3.	Occupation	Nominal: Job position
4.	Unpleasant childhood experience	Nominal: Yes or no
5.	Depression	Nominal: Yes or no based on the scores obtained

Table 2. Attributes based on BDI [14]

No.	Attribute	Type
1.	Sadness	Numeric: {0,1,2,3}
2.	Pessimism	Numeric: {0,1,2,3}
3.	Past failure	Numeric: {0,1,2,3}
4.	Loss of pleasure	Numeric: {0,1,2,3}
5.	Guilt feelings	Numeric: {0,1,2,3}
6.	Punishment feelings	Numeric: {0,1,2,3}
7.	Self-dislike	Numeric: {0,1,2,3}
8.	Self-criticalness	Numeric: {0,1,2,3}
9.	Suicidal thoughts or wishes	Numeric: {0,1,2,3}
10.	Crying	Numeric: {0,1,2,3}
11.	Agitation	Numeric: {0,1,2,3}
12.	Loss of interest	Numeric: {0,1,2,3}
13.	Indecisiveness	Numeric: {0,1,2,3}
14.	Worthlessness	Numeric: {0,1,2,3}
15.	Loss of energy	Numeric: {0,1,2,3}
16.	Changes in sleeping pattern	Numeric: {0,1,2,3}
17.	Irritability	Numeric: {0,1,2,3}
18.	Changes in appetite	Numeric: {0,1,2,3}
19.	Concentration difficulty	Numeric: {0,1,2,3}
20.	Tiredness or fatigue	Numeric: {0,1,2,3}
21.	Loss of interest in sex	Numeric: {0,1,2,3}

Inventory (BDI) [3] (Table 2), which is one of the most widely used psychometric tests measuring severity of depression. The attributes and their descriptions are summarized and listed on the tables below.

5 Results and Evaluation

This section shows the experimental and the analysis of the result. In this study the performance evaluators of four classification techniques using three approaches were compared and analyzed over anxiety disorders dataset obtained from the online-based survey

Table 3. Classification performances for anxiety disorders dataset

Authors	Classifier	Approach	No. of features	Accuracy (%)	Sensitivity (%)	Specificity (%)
Husain et al. 2016	Random forest	Random forest	3	94.64	45.45	100
Sribala M. & M. 2015	Neural network	Neural network (with sensitive analysis)	12	–	96.42	–
Our work	Naive Bayes	Naive Bayes	26	91.76	90.91	100
		Naive Bayes+PSO	14	93.41	92.73	100
		Naive Bayes+PSO+FRS	4	93.96	95.76	76.47
	Support Vector Machine (SVM)	SVM	26	96.70	98.79	76.47
		SVM+PSO	14	96.15	98.18	76.47
		SVM+PSO+FRS	4	92.86	99.39	29.41
	K-Nearest Neighbor (KNN)	KKN	26	96.15	98.18	76.47
		KNN+PSO	14	93.41	96.97	58.82
		KNN+PSO+FRS	4	91.76	98.79	23.53
	Decision tree	Decision tree	26	91.21	94.55	58.82
		Decision Tree+PSO	14	90.11	94.55	47.06
		Decision Tree+PSO+FRS	4	93.96	100	35.29

Table 3 shows the overall performances of classifier model on the anxiety disorders dataset. It shows the number of features selected by proposed approach, classification accuracy, sensitivity and specificity. The features are selected by proposed approach either by PSO or feature selection with PSO. The selected features are then applied on the four classifiers which are Naïve Bayes, SVM, KNN and decision tree. The four classifiers are then tested on anxiety disorders dataset using the features as identified by the proposed approach. There are originally 26 features available on the dataset. By applying PSO in each classifier, the features are reduced to 14 features for further data mining. The features are even further reduced to 4 features in the model using PSO and Fuzzy Rough Set feature selection method in classifier. The final four features selected are occupation, pessimism, suicidal thoughts or wishes and depression.

The features are selected by proposed approach either by PSO or feature selection with PSO. The selected features are then applied on the four classifiers which are Naïve Bayes, SVM, KNN and decision tree. The four classifiers are then tested on anxiety disorders dataset using the features as identified by the proposed approach. There are originally 26 features available on the dataset.

By applying PSO in each classifier, the features are reduced to 14 features for further data mining. The features are even further reduced to 4 features in the model using PSO and Fuzzy Rough Set feature selection method in classifier. The final four features selected are occupation, pessimism, suicidal thoughts or wishes and depression.

5.1 Accuracy

Based on the results in Table 3, the overall accuracy for all models are above 90 %. The results also show that there is an improvement in classification accuracy of two classifiers on anxiety disorders dataset as a result of decreasing the number of features. The two classifiers are Naïve Bayes and decision tree. Both classifiers are able to enhance its data classification accuracy when reducing the number of features. However for SVM and KNN classifiers, decreasing the number of features did not improve its accuracy. Instead, its accuracy decreases when the number of features decreases. This shows that performance of SVM and KNN classifiers has not enhanced accuracy with the lesser number of features by feature selection and PSO. Among all classifiers, the model which gives the highest accuracy is the model using only SVM classifier. Without implementation of feature selection or PSO, the SVM model obtains the accuracy of 96.70 %, which is even higher than the classifier model using PSO and feature selection.

5.2 Sensitivity

From the results in Table 3, it can be observed that classification sensitivity improves when there is lesser number of features for all classifiers on anxiety disorders dataset. Minimize the number of features used for classification is found to be able to increase sensitivity of the model. The overall sensitivity of all model are above 90 %. The lowest sensitivity is 90.91 % for the model using only Naïve Bayes classifier. While the model that gives the highest sensitivity of 100 % is the model applying PSO and feature selection with decision tree classifier.

5.3 Specificity

Based on results obtained from Table 3, there is a drastic difference of specificity between the models. Unlike accuracy and sensitivity, specificity of all models is between the ranges of 20 % to 100 %. The lowest specificity is 23.53 % for the model using PSO and feature selection with KNN classifier. Meanwhile, both models of Naïve Bayes classifier with and without applying PSO give the highest specificity of 100 %. The specificity decreases for all classifiers that implementing features selection and PSO. When the number of features decreases, it affects the performance by lowering its specificity. Therefore, it is found that reducing the number of features does not improve specificity of all four classifiers for anxiety disorders dataset.

6 Conclusion

The proposed approaches have been experimented with generalized anxiety disorders data. The experimental results clearly show a perceivable improvement in classification accuracy for all classifiers except for SVM and KNN classifiers. For Naïve Bayes and decision tree classifiers, the reduction in features improves its performances in accuracy aspect. Besides, based on the performance evaluation on sensitivity, there is an improvement in sensitivity for all four classifiers when the number of features selected is lesser. For the performance evaluation on specificity, the results show tremendously low specificity for all classifiers model implementing PSO and feature selection. Instead, all classifier models that using all features in dataset apparently give a better result in specificity. It is found that the implementation of PSO on generalized anxiety disorders dataset has different results in performance depending on the selection of classification technique.

7 Future Work

This study mainly focuses on PSO algorithm and Fuzzy Rough Set as the feature selection method. Therefore the future work will focus on applying different evolutionary algorithm such as Ant Colony Optimization (ACO). Besides, other feature selection methods such as Correlation-based feature selection (Cfs), Information Gain (IG), Gain Ratio and Chi-Squared (CHI) can also be used to test the model performance of the proposed approach on generalized anxiety disorder data.

Acknowledgments. We would like to express our gratitude to Universiti Sains Malaysia (USM) for supporting this research.

References

1. Ajith, A., Guo, H., Liu, H.: Swarm intelligence: foundations, perspectives and applications. In: Nedjah, N., de Macedo Mourelle, L. (eds.) Swarm Intelligent Systems. SCI, vol. 26, pp. 3–25. Springer, Heidelberg (2006)
2. Engelbrecht, A.P.: Fundamentals of Computational Swarm Intelligence. Wiley, New York (2006)
3. Alam, S., Dobbie, G., Koh, Y.S., Riddle, P., Rehman, S.U.: Research on particle swarm optimization based clustering: a systematic review of literature and techniques. Swarm Evol. Comput. **17**, 1–13 (2014)
4. Hong, H., Li, J., Plank, A., Wang, H., Daggard, G.: A comparative study of classification methods for microarray data analysis. In: Proceedings of Fifth Australasian Data Mining Conference on Australasian Data Mining Conference (AusDM2006), pp. 33–37 (2006)
5. Potter, R.: Comparison of classification algorithms applied to breast cancer diagnosis and prognosis. In: 7th Industrial Conference on advances in data mining (ICDM 2007), pp. 40–49. Leipzig, Germany (2007)

6. Lin, S.-W., Ying, K.-C., Chen, S.-C., Lee, Z.-J.: Particle swarm optimization for parameter determination and feature selection of support vector machines. Expert Syst. Appl. **35**(4), 1817–1824 (2008)
7. Chakraborty, B.: Feature subset selection by particle swarm optimization with fuzzy fitness function. In: 3rd International Conference on Intelligent System and Knowledge Engineering (ISKE 2008). IEEE (2008)
8. Ming, H.: A rough set based hybrid method to feature selection. In: International Symposium on Knowledge Acquisition and Modelling (KAM 2008). IEEE (2008)
9. Gao, H.-H., Yang, H.-H., Wang, X.Y.: Ant colony optimization based network intrusion feature selection and detection. In: International Conference on Machine Learning and Cybernetic 2005 (ICMLC 2005). IEEE (2005)
10. Pawlak, Z.: Rough Sets: Theoretical Aspects of Reasoning about Data. Springer, Netherlands (1991)
11. Eberhart, R., Kennedy, J.: A new optimizer using particle swarm theory. In: Proceedings of the Sixth International Symposium on Micro Machine and Human Science (MHS 1995). IEEE (1995)
12. Sousa, T., Silva, A., Neve, A.: Particle swarm based data mining algorithms for classification tasks. Parallel Nat. Inspired Comput. Paradigms Appl. **30**(5), 767–783 (2004)
13. Holden, N., Freitas, A.A.: A hybrid particle swarm/ant colony algorithm for the classification of hierarchical biological data. In: IEEE Swarm Intelligence Symposium (SIS 2005). IEEE (2005)
14. American Psychiatric Association: Diagnostic and Statistical Manual of Mental Disorders. American Psychiatric Publishing, Arlington (2013)

Quality Improvement of Vietnamese HMM-Based Speech Synthesis System Based on Decomposition of Naturalness and Intelligibility Using Non-negative Matrix Factorization

Anh-Tuan Dinh[1(✉)], Thanh-Son Phan[2], and Masato Akagi[1]

[1] Graduate School of Advanced Science and Technology,
Japan Advanced Institute of Science and Technology,
1-1 Asahidai, Nomi, Ishikawa, Japan
{tuan.dinh,akagi}@jaist.ac.jp
[2] Faculty of Information Technology, Telecommunications University,
101 Mai Xuan Thuong, Nha Trang, Khanh Hoa, Vietnam
ptson@tcu.edu.vn

Abstract. Hidden Markov model (HMM)-based synthesized speech is intelligible but not natural especially under limited data condition. The goal of this study is to improve naturalness without violating acceptable intelligibility by decomposing the naturalness and intelligibility of synthesized speech using a novel asymmetric bilinear model involving non-negative matrix factorization (NMF). Subjective evaluations carried out on Vietnamese data confirmed that the achieved synthesis quality is higher than other methods under limited data condition. Since F0 contour is important for naturalness and intelligibility, especially in Vietnamese. Proposed method is capable of modifying over-smoothed F0 contour without destroying tonal information.

Keywords: Hidden markov model · Non-negative matrix factorization · Naturalness-intelligibility decomposition

1 Introduction

Although Vietnamese is spoken by about 100 million people, there is no huge public speech-corpus with labelling for Vietnamese. In other word, Vietnamese is an under-resourced language. Therefore, to build a Vietnamese text-to-speech system, we need a speech-synthesis technique which can perform well under limited data condition.

Hidden Markove model (HMM)-based speech synthesis is a state-of-the-art method due to its flexibility and compact footprint [1]. Therefore, building Vietnamese HMM-based speech synthesis system has recently achieved promising results [11,12]. The HMM can model not only the statistical distribution of

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_53

speech parameters but also their rate of change. As a result, synthesized speech is intelligible but not natural due to statistical averaging or over-smoothing of speech parameters (e.g., speech spectra and F0) especially under limited data conditions. There have been several attempts to overcome the over-smoothing of speech parameters. Thus, it is a challenge to improve naturalness without violating intelligibility especially for Vietnamese synthesized speech.

One approach is using objective evaluations of this effect such as global variance (GV) [2], and modulation-spectrum [3], integrating them into the parameter generation phase to obtain better speech parameter values. Context-dependent models are usually trained for objective evaluations. However, under limited data conditions, there is not enough data to train the context-dependent models for all possible contexts. Another possible approach is to reduce the gap between the speech parameters of natural and that of synthetic speech by learning the acoustic differences directly from the data. If we consider natural speech as target speech and synthesized speech of the same sentence, same speaker characteristics as source speech, voice conversion techniques [4,5] can be used as mapping from synthetic speech to natural speech. Since quality improvement is independent from synthesizers, we can improve the naturalness of current speech synthesizers. Thus, a voice-conversion approach is used to improve naturalness.

With the majority of previous voice-conversion approaches, all spectra and F0 are modified to improve naturalness. However, applying these approaches often negatively affect intelligibility e.g., destroying tonal information. This drastically affects intelligibility of tonal languages such as Vietnamese and Chinese. To improve naturalness without violating intelligibility, an asymmetric bilinear model [6] was introduced to decompose naturalness and intelligibility. Popa et al. [7] used an asymmetric bilinear model to decompose a speech parameter vector into speaker information and phonetic information using singular value decomposition (SVD). From this idea, a speech parameter vector y can be represented as a combination of the naturalness factor $\mathbf{A}$ and intelligibility factor $\mathbf{b}^c$ of an intelligibility class c:

$$\mathbf{y} = \mathbf{A}\mathbf{b}^c \tag{1}$$

In the representation of an asymmetric bilinear model, naturalness can be modified, whilst intelligibility can be preserved. One problem with applying an asymmetric bilinear model is finding an efficient constraint to decompose naturalness and intelligibility from speech spectra and F0 contour. Although an asymmetric bilinear model using SVD is an excellent approach, SVD allows negative combinations of intelligibility and naturalness. Since combinations indicate unrealistic subtractions of intelligibility (or naturalness), negative combinations are unnecessary. To avoid subtractive combinations, we propose a method that uses our novel asymmetric bilinear model involving NMF.

An overview of modifying HMM-based synthesized speech using asymmetric bilinear model involving NMF is shown in Fig. 1. Our proposed method is used in spectra modification and F0 modification.

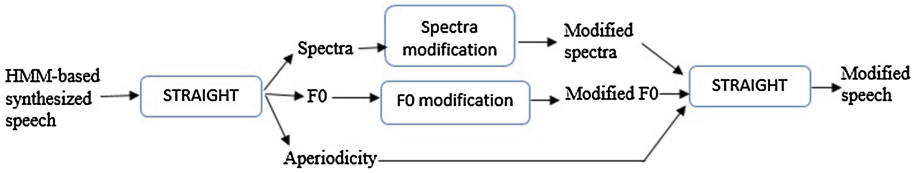


Fig. 1. Scheme of modifying Vietnamese HMM-based synthesized speech

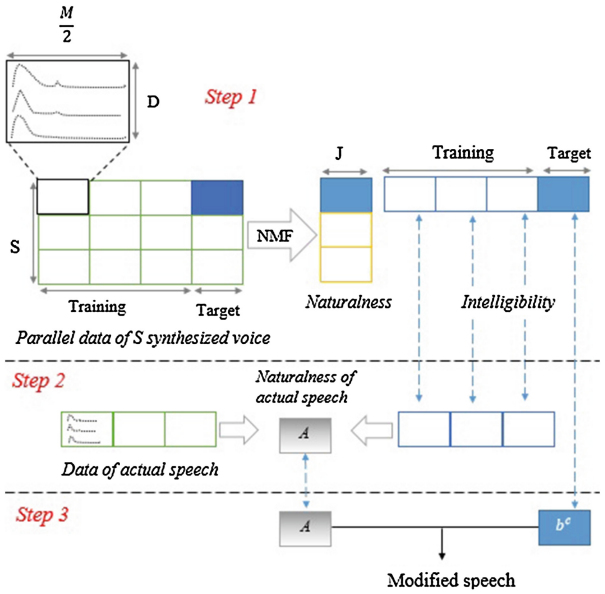


Fig. 2. Scheme of applying our proposed method for improving synthesized speech-spectra; N training sentences and 1 target sentence

2 Spectra Modification

In the section, we describe the process of applying an asymmetric bilinear model for improving synthesized speech-spectra. This speech spectra is represented by Mel-cepstral coefficient (MCC) ($\gamma = 0$, $\alpha = 0.42$ for 16 kHz speech) [9]. We used the modulation-spectrum of MCC sequences $\mathbf{c}_k = [c_{1k}, c_{2k}, ..., c_{Dk}]^\top$, $k = 1, 2, ..., T$, in which D is the order of cepstral coefficients and T is the number of frames, to determine the over-smoothing effect in both the time and frequency domains of speech spectra [10]. Short-term spectral analysis of a speech utterance yields a matrix $R = [\mathbf{c}_1, \mathbf{c}_2, ..., \mathbf{c}_T]$ of size $D \times T$. The time trajectory of cepstral coefficient d is defined as $\mathbf{r}_d = [c_{d1}, c_{d2}, ..., c_{dT}]$, $d = 1, 2, ..., D$. The modulation-spectrum of trajectory r_d is defined as:

$$M(d, f) = |FT[\mathbf{r}_d]|, \tag{2}$$

where f is the modulation frequency bin, defined by the number of points in the Fourier transform (FT). The number of points in the FT must be greater than the maximum number of frames T of an utterance. The modulation-spectrum of each utterance is calculated for each coefficient. Using an asymmetric bilinear model, the modulation-spectrum of synthetic speech-spectra is modified to be closer to the modulation characteristics of natural one.

The process of improving speech spectra consists of three major steps as shown in Fig. 2:

1. Decomposition of naturalness and intelligibility of synthesized voices.
2. Obtaining naturalness of actual speech.
3. Reconstructing modified speech with intelligibility of synthesized voice and naturalness of actual speech.

2.1 Decomposition of Naturalness and Intelligibility of Synthesized Voices

The goal with step 1 is to obtain acceptable intelligibility from parallel data of synthesized voices to preserve the intelligibility. The naturalness and intelligibility factors were factorized from the data using NMF.

Stacking Parallel Data of S Synthesized Voices: We first prepared parallel data of a number of S HTS voices, as shown in Fig. 3. In the parallel data, the variation in different HTS voices' quality is presented in columns and that in phonetic information of different sentences is presented in rows. Phonetic information is assumed to be intelligible. To build this parallel data of a number of S HTS voices, the modulation-spectrum of the MCC sequence from N sentences were stacked horizontally when the parallel data of a number of S HTS voices is decomposed into two components: naturalness and intelligibility, as shown in Fig. 3, where M denotes the number of FT points for modulation-spectrum, D is the MCC order, N is the number of sentences, S is the number of Vietnamese HTSs [11,12] (S $\geq$ 2), and J is the number of model dimensions determined as $J = S \times D$ [7].

2.2 Natural Improvement of HMM-Based Synthesized Speech

In step 2, the naturalness of actual speech $\mathbf{A}$ was obtained using a small amount of actual speech y and corresponding intelligibility set C obtained from step 1. We derived the desired naturalness $\mathbf{A}$ by minimizing the total squared error over actual speech data,

$$E = \sum_{c \in C} ||\mathbf{y} - \mathbf{A}\mathbf{b}^c||^2 \tag{3}$$

In Eq. 3, intelligibility vectors $\mathbf{b^c}$ were learned from step 1. The desired naturalness $\mathbf{A}$ can be found by solving the linear system

$$\frac{\partial E}{\partial \mathbf{A}} = 0 \tag{4}$$

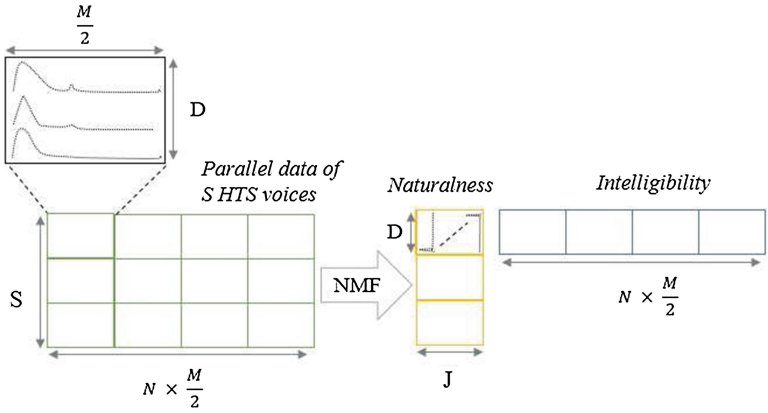


Fig. 3. Our asymmetric bilinear model using NMF; S: same sentence and different HTSs

In step 3, the naturalness of actual speech **A** and intelligibility of synthesized speech were combined to obtained an improved version of synthesized speech.

3 The F0 Modification

With Vietnamese, intonation managed by F0 contour is important for perceiving naturalness [12]. Moreover, tonal information determined by F0 contour is also important for intelligibility. In order to improve quality of Vietnamese HMM-based speech synthesis, modifying over-smoothed F0 contour is necessary.

3.1 Tonal Information in Vietnamese

The Vietnamese is a monosyllabic and tonal language with six tones. According to [13], a syllable structure can be described as in Table 1. A syllable consists of initial and final part. The final part can be further divided into onset, nucleus, and coda. The onset and coda parts are optional. Six Vietnamese tones are: level, falling, broken, curve, rising, and drop.

Table 1. Structure of Vietnamese syllable

Tone			
Initial	Final		
	[Onset]	Nucleus	[Coda]

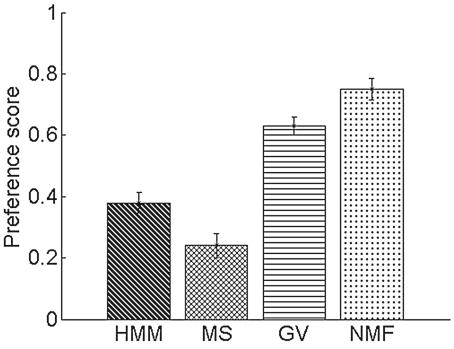


Fig. 4. Preference scores with 95 % confidence interval

3.2 Quality Improvement of F0 Contour

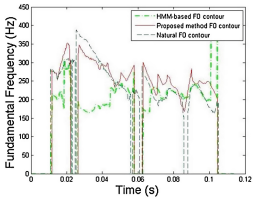
The process of applying our asymmetric bilinear model involving NMF for modifying synthesized F0 contour is the same as that of applying the method for modifying synthesized speech-spectra. In Step 1, continuous F0 contour, interpolated in unvoiced regions, from N sentences were stacked horizontally to build parallel data of a number of S HTSs. Then, the parallel data is decomposed into two components: naturalness and intelligibility. Intelligibility component, which consists of phonetic and tonal information, is preserved when we modify naturalness component.

In Step 2, naturalness component of natural F0 contour is derived from a small data of actual speech. By combining naturalness component of natural F0 contour and intelligibility of synthesized F0 contour in Step 3, we obtain modified F0 contour.

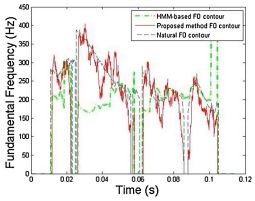


(a) Tone intelligibility test (b) MRT

Fig. 5. Experimental results with 95 % confidence interval



(a) Use F0 values as input of proposed method



(b) Use modulation-spectrum of F0 contour as input of proposed method

Fig. 6. The F0 contours of proposed method involving NMF, HMM-based synthesized speech, and natural speech

4 Evaluation and Discussion

We evaluated the naturalness and intelligibility of the proposed method using a preference test, a Vietnamese-tone intelligibility test, and a modified rhythm test (MRT) under limited data conditions.

In the preference test, the proposed method was compared with other improvement methods such as those involving GV [2], and modulation-spectrum [4]. Two HMM-based synthesized voices ($S = 2$) were trained using 2 Vietnamese datasets (DEMEN567 and FEMALE1). The DEMEN567 was called TTSCorpus in [14]. This corpus has 567 utterances with sampling rate of 11025 Hz. The FEMALE1 includes 567 utterances with sampling rate of 16000 Hz. This dataset was used in [12] to show an improved naturalness of Vietnamese HMM-based synthesized speech by adding prosodic information to label files. To simulate the limited data conditions, 500 utterances in each data set were used to train 2 different HMM-based synthesizers. Both of the two datasets have a phoneme coverage of one hundred percentage. Fifteen sentences were chosen as testing sentences. These sentences were synthesized using synthesizer trained from FEMALE1 dataset as baseline. We applied proposed method and those involving GV, and modulation-spectrum to improve the quality of the samples under limited data conditions. A number of 500 natural utterances were used for training both methods involving GV and modulation-spectrum. To improve spectral parameters of the synthesized speech using our proposed method involving NMF, we synthesized 4 training sentences using $S = 2$ synthesizers trained from DEMEN567 and FEMALE1 datasets (in total 8 training utterances) in step 1. We also synthesized fifteen testing sentences using $S = 2$ synthesizers trained from DEMEN567 and FEMALE1 datasets (in total 30 testing utterances). With each testing sentence, a parallel data of spectral parameters was formed using spectral parameters of 8 training utterances and 2 testing utterances with $S = 2$ and $N = 5$ sentences. The objective is to modify the testing utterance generated by synthesizer trained from FEMALE1 dataset. In step 2, original utterances of four training sentences from FEMALE1 dataset was also used to derive naturalness of actual speech. We combine the naturalness of actual speech with intelligibility of testing sentence obtained from step 1 to obtain modified

spectral parameter for the testing sentence. The process of modifying F0 contour is the same as that of modifying spectral parameter. All HMM-based synthesized utterances were aligned with their original human-speech using the guide of label files from FEMALE1 dataset. The STRAIGHT vocoder [16] was used to analyze the speech. The frame-shift was 5 ms and the frame-length was 10 ms. It decomposes speech into a spectral envelope, F0, and aperiodicity. Linear interpolation was used to generate F0 values in unvoiced regions. The STRAIGHT-based spectral parameters are further encoded into MCC. The cepstral order was 49 and the MCC sequences were transformed into the modulation-spectrum using FT; $M = 4096$. Eleven individuals (six northern-, and five southern-Vietnamese people) listened to 180 pairs of utterances. The participants are graduate students with normal hearing ability. They listened to each pair only once, then compared the naturalness of utterances on a two-point scale, i.e., 1 (A is more natural), and -1 (B is more natural). Natural speech was defined as actually human-speech.

Figure 4 shows that preference score of proposed method (denoted as NMF) is the best under limited data condition. Since there was not enough data to train context-dependent models considering tonal-information, the method involving GV (denoted as GV) did not perform well. Since the method involving modulation-spectrum (denoted as MS) does not consider tonal-information, it destroyed tonal-information and generated unnatural utterances. Therefore, its preference score is the worst. At the end of the experiments, participants were asked what factors contribute to their decisions. All participants agreed that speech with buzzing sound and meaningless speech is not natural.

In the tonal-intelligibility test, we evaluate the tonal-intelligibility of synthesized speech after applying proposed method. We prepared 50 syllables synthesized by Vietnamese HMM-based synthesized speech [12]. The syllables were randomly selected from 67 remaining sentences in FEMALE1 dataset. Different methods involving NMF, GV, and modulation-spectrum were used to improve naturalness of the syllables. All configurations were the same as previous experiment. Twelve individuals (seven northern-, and five southern-Vietnamese people) listened to 250 syllables from HMM-based synthesized speech, proposed method involving NMF, other methods involving GV and modulation-spectrum, and human-speech. The 50 syllables spoken by human were used as reference. The participants are graduate students with normal hearing ability. Participants listened to each syllable only once, and selected the most likely syllable they heard among a group of syllables bearing different tones (e.g., la, là, lá, lả, lã, lạ). For each syllable, we put "not like all above syllables" in answer list. The correctness for each method was obtained by pooling responses for the method, tallying the total number of correctly selected syllables, dividing by the total number of listened-syllables, and multiplying by 100.

Figure 5a shows that the correctness of proposed method (denoted as NMF) is equal to that of HMM-based synthesized speech (denoted as HMM). It indicates that the tone-intelligibility of synthesized speech was preserved by proposed method. In contrast, the tone-intelligibility of synthesized speech was destroyed by the method involving modulation spectrum (denoted as MS).

In the modified rhyme test (MRT), we evaluate the intelligibility of synthesized speech after applying proposed method. We prepared 60 syllables synthesized by Vietnamese HMM-based synthesized speech [12]. The syllables were randomly selected from 67 remaining sentences in FEMALE1 dataset. Different methods involving NMF, GV, and modulation-spectrum were used to improve naturalness of the syllables. All configurations were the same as previous experiment. Eleven individuals (six northern-, and five southern-Vietnamese people) listened to 300 syllables from HMM-based synthesized speech, proposed method involving NMF, other methods involving GV and modulation-spectrum, and human-speech. The participants are graduate students with normal hearing ability. Participants listened to each syllable only once, and selected the most likely syllable they heard among a group of syllables bearing different final part (e.g., la, lan, lanh, lang). For each syllable, we put "not like all above syllables" in answer list.

Figure 5b shows that the correctness of proposed method (denoted as NMF) is equal to that of HMM-based synthesized speech (denoted as HMM). It indicates that the intelligibility of synthesized speech was preserved by proposed method. In contrast, the intelligibility of synthesized speech was destroyed by the method involving modulation-spectrum (denoted as MS). Natural speech is denoted as NAT.

When we apply proposed method involving NMF to modify F0 contour of synthesized speech, both HMM-based F0 values and modulation-spectrum of continuous HMM-based F0 contour were considered as input of our proposed method. In [4], a post-filter was applied to modify modulation-spectrum of continuous HMM-based F0 contour. However, applying our proposed method on modulation-spectrum of HMM-based F0 contour did not yield a good result. Unexpected fluctuations were added to obtained F0 contour by proposed method as in Fig. 6b. But, applying our proposed method on HMM-based F0 contours yielded better result as in Fig. 6a. The reason may be F0 contour is smoother than speech spectra. Improving modulation-spectrum of F0 contour means improving F0 contour's fine-structure which generate unexpected fluctuation for obtained F0 contour from prposed method.

5 Conclusion

We proposed a novel asymmetric bilinear model using NMF to decompose the naturalness and intelligibility of Vietnamese HMM-based synthesized speech. The proposed method proved to be efficient in improving naturalness without violating the intelligibility of synthesized speech, especially under limited data condition. The proposed method can model and modify F0 contour of Vietnamese. Experimental results demonstrated its superiority to other methods under a limited data condition. Our method provides a new way to control naturalness of synthesized speech under limited data conditions.

Acknowledgement. This study was supported by the Grant-in-Aid for Scientic Research (A) (No. 25240026), SECOM Science and Technology Foundation and the JSPS A3 Foresight program.

References

1. Zen, H., Tokuda, K., Black, W.: Statistical parametric speech synthesis. Speech Comm. **51**(11), 1039–1064 (2009)
2. Toda, T., Tokuda, K.: A speech parameter generation algorithm considering global variance for HMM-based speech synthesis. IEICE Trans. **E90–D**(05), 816–824 (2007)
3. Takamichi, S., Toda, T., Black, A., Nakamura, S.: Parameter generation algorithm considering modulation spectrum for HMM-based speech synthesis. In: Proceedings of ICASSP, pp. 4210–4214 (2015)
4. Takamichi, S., Toda, T., Neubig, G., Nakamura, S.: A post-filter to modify the modulation spectrum in HMM-based speech synthesis. In: Proceedings of ICASSP, pp. 290–294 (2014)
5. Chen, L.H., Raitio, T., Valentini-Botinhao, C., Yamagishi, J., Ling, Z.H.: DNN-based stochastic postfilter for HMM-based speech synthesis. In: Proceedings of Interspeech, pp. 1954–1958 (2014)
6. Tenenbaum, J., Freeman, W.: Separating style and content with bilinear models. Neural Comput. **12**, 1247–1283 (2000)
7. Popa, V., Nurminen, J., Gabbouj, M.: A novel technique for voice conversion based on style and content decomposition with bilinear models. In: Proceedings of Interspeech, pp. 2655–2658 (2009)
8. Stylianou, Y., Cappe, O., Moulines, E.: Continuous probabilistic transform for voice conversion. IEEE Trans. Audio, Speech, Lang. Process. **6**, 131–142 (1998)
9. Tokuda, K., Masuko, T., Imai, S.: Mel-generalized cepstral analysis - a unified approach to speech spectral estimation. In: Proceedings of ICSLP, pp. 1043–1046 (1994)
10. Dinh-Anh, T., Morikawa, D., Akagi, M.: Study on quality improvement of HMM-based synthesized voices using asymmetric bilinear model. J. Sig. Process. **20**(4), 205–208 (2016)
11. Vu, T.T., Luong, M.C., Nakamura, S.: An HMM-based vietnamese speech synthesis system. In: Proceedings of Oriental COCOSDA, pp. 116–121 (2009)
12. Phan, T.S., Duong, T.C., Dinh, A.T., Vu, T.T., Luong, M.C.: Improvement of naturalness for an HMM-based Vietnamese speech synthesis using the prosodic information. In: Proceedings of RIVF, pp. 276–281 (2013)
13. Doan, T.T.: Ngữ âm tiếng Việt (Vietnamese Phonetics), pp. 99–148. Hanoi National University Publishing House (1999)
14. Mai, L.C., Duc, D.N.: Design of Vietnamese speech corpus and current status. In: Proceedings of ISCSLP 2006, pp. 748–758 (2006)
15. Scheffe, H.: An analysis of variance for paired comparisons. J. Am. Stat. Assoc. **37**, 381–400 (1952)
16. Kawahara, H., Masuda-Katsue, I., de Cheveigne, M.: Restructuring speech representations using a pitch-adaptive time-frequency smoothing and a instantaneous frequency-based F0 extraction: Possible role of a repetitive structure in sounds. Speech Comm. **27**, 187–207 (1999)

Reducing Middle Nodes Mapping Algorithm for Energy Efficiency in Network Virtualization

Tran Manh Nam(✉), Nguyen Van Huynh, and Nguyen Huu Thanh

School of Electronics and Telecommunications,
Hanoi University of Science and Technology, Hanoi, Vietnam
nam.tranmanh-set@hust.edu.vn
http://set.hust.edu.vn

Abstract. For the future of the Internet, Network Virtualization and Software-Defined Networking (SDN) are recognized as key technologies. They reshape computing and network architectures, provide a number of advantages including centralized management, scalability and resource optimization. Along with the Internet's development, the networking devices such as routers and switches are notable part of the large energy consumption of ITC. The design of high performance and energy-efficient network by using virtualization and SDN becomes an importance issue. In this paper we heuristic present energy-efficient algorithms for virtual network embedding. The experimental results show the remarkable energy-saving level of their approaches while maintaining the acceptance ratio.

Keywords: Energy efficiency · Software-Defined Networking · Network virtualization · Middle node

1 Introduction

Nowadays, the Internet services are quickly growing in the size and number of services such as Youtube, Facebook, Dropbox...etc. Therefore, carbon emissions from the Internet sector such as Data Centers, ISPs are estimated to keep rising significantly over the coming years from 0.53 billion tons (Gt) carbon dioxide equivalent (CO2e) in 2002 to 1.43 Gt CO2e in 2020 under BAU growth [1]. Consequently, the power-saving ITC solutions have attracted much attention from research community. In the ICT sector, the novel business model called *Infrastructure as a Service* (IaaS) that based on *Network Virtualization* (NV) technology is become promising model. With network virtualization, multiple heterogeneous virtual network, the main entity of NV, can run on the same physical network and share the resources with each other. In IaaS, the ISP can be divided into two different layers, namely *Infrastructure Provider* (IP) that owns such physical network resources and *Service Provider* (SP) who deploys network protocols and offers end-to-end services. The SP layer contains three sub-layers: (1) *Virtual network Provider* that collects and manages physical resources from

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_54

several Infrastructure Providers; (2) the *Virtual Network Operator* who makes use of underlying physical resources to create virtual networks; and (3) *Service Provider* who creates and offers services through virtual networks [NV-Survey-botero].

Network virtualization allows many virtual networks allocated in the same physical network. The most important problem of NV is *virtual network embedding (VNE)* problem that maps separate *virtual network requests (VNRs)* onto the same substrate network to satisfy the objective of the infrastructure provider. The VNE problem can be divided to two sub-problems, namely *virtual node mapping (VNoM)* that maps the virtual nodes of a VNR onto the substrate node and *virtual link mappling (VLiM)* that maps the virtual links among virtual nodes onto the substrate link. In NV, one virtual node can be mapped to only one substrate node, while a virtual link can be represented by a path which stays on group of consecutive physical links in substrate network. We consider that a *middle node* is an intermediate node in the subtrate network path mapped to a virtual link. For energy efficiency of network virtualization, we can see that middle nodes constantly consume power although they are transparent to the customer. So that the VNE algorithm should focus onto minimizing the number of middle nodes, consequently, we can turn the unused component of the physical network off for saving energy. In this paper, we propose a reducing middle-node VNE algorithm (RmnE), our contributions are described as follows:

- Proposing two energy-efficient embedding algorithms including greedy capacity algorithm and reducing middle node (RMN) algorithm. The RMN aims to reduce middle nodes in substrate network.
- Combining a *idle logic* and a *power scaling* energy-efficient strategies with mapping algorithm. The experimental results show that our proposed algorithm increases the power-saving level while maintains the its acceptance ratio in comparison with some existing algorithms.

This article is organized as follows: Sect. 2 describes the background and Sect. 3 presents the problem statement and formulation; Our energy-aware network virtualization system is described in Sect. 4; finally, conclusions are drawn in Sect. 5.

2 Related Work

2.1 Network Energy-Efficient Approaches

There are two main reasons that make the *green ICT* term to become a most important topic. Firstly, the environmental aspect, which is related to the reduction of CO_2 emission. Secondly, the economical aspect, which is related to the costs sustained by the service providers. From general point of view, the largest part of undertaken approaches is founded on few basic concepts which can be classified as follows [9]: *re-engineering; dynamic adaptation; and sleeping/standby.* Among these approaches, *dynamic adaptation* has been attracted

much attention from research communities. This approach is generally founded on two main power-aware capabilities, namely *power scaling* and *idle logic*. Idle logic approach allows reducing power consumption by quickly turning off subnetwork when there is no traffic demand and by rapidly waking up them when necessary. On the other hand, power scaling allows dynamically change the working speed of network devices to adapt to the network utilization. Although these aforementioned approaches can be found in the literature, a recently-rising question is: can we combine them together and how does it work in network virtualization. In this paper, we deploy and evaluate the performance of power scaling and idle logic in terms of the power saving level under different load in network virtualization environment.

2.2 Network Hypervisor

Network Virtualization concept is realizing by using network *hypervisor-like layer* that sits between the underlying physical network and the software that controls it. By using Software-Defined Networking technology, we can realize the NV concept and create many virtual networks on top of substrate network. There are some NV platforms using the SDN protocol stack including FlowVisor [2,8] etc. There layers are used as a middle layer among the controllers and switches - provides fast, easy and transparent decomposition of a given network into virtual slices and delegates slice messages to the controller. In [3], Pham et al. successfully extended the gigabit Net-FPGA switch to be a energy-aware switch which supports Openflow protocol and works well with SDN controller. So that, the deployment of energy-aware network virtualization platform is feasible.

2.3 Virtual Network Embedding Algorithm

Network virtualization allows many virtual networks allocated in the same physical network. Currently, the VNE problem has been extensively investigated in the literature for different purposes, such as improving acceptance ratio of VNRs or maximizing the revenue by accommodating more VNRs in the substrate network. In [5] Houidi et al. use the VNE ILP formulation to seek for the minimization of the embedding cost and the maximization of the acceptance ratio. For the energy-efficient VNE, there are some algorithms that have been proposed. Botero et al. proposed exact formulations - Energy Aware VNE (VNE-EA) [4] and Energy Aware VNE and Load Balancing (VNE-EA-LB) [6] that used a *Mixed Integer Program* (MIP) to optimally solve the embedding problem. Integer Linear Programs (ILPs) are in many practical situations, NP-hard, and therefore, the implementation of these models is not scalable in large networks. On the other hand, these exact algorithms also increase the delay time of the VNRs mapping, so that it's necessary a heuristic embedding algorithm that achieves near-optimal energy-saving level while maintaining the acceptable processing times.

3 Problem Modeling

3.1 VNE Modeling

In this work, a substrate (physical) network is described as weighted undirected graph $G^s = (N^s, L^s)$ where N^s and L^s are set of substrate nodes and links, respectively. Each substrate node $n^s \in N^s$ represents a switch with the CPU capability $C(n^s)$. Each substrate link $l^s(x^s, y^s) \in L^s$, interconnects between substrate nodes x^s and y^s, and has bandwidth $B(l^s)$. We denote P^s as a set of all substrate paths in the network while $P^s(s^s, d^s)$ is a set of all available paths from source node s^s to destination node d^s. Similarly, we denote the virtual network request i^{th} as $VNR_i = G_i^v(N_i^v, L_i^v)$, where N_i^v is a set of virtual nodes and L_i^v is a set of virtual links. A virtual node $n^v \in N_i^v$ requires the physical resource $C(n^v)$ while virtual link $l^v \in L_i^v$ request bandwidth requirement $B(l^v)$. We define $aC(n^s)$ and $aB(l^s)$ as the available CPU of substrate node $n^s \in N^s$ and available bandwidth of substrate link $l^s \in L^s$, respectively (Eqs. 1 and 2).

$$aC(n^s) = C(n^s) - \sum_{\forall n^v \to n^s} C(n^v) \tag{1}$$

$$aB(l^s) = B(l^s) - \sum_{\forall l^v \to l^s} B(l^v) \tag{2}$$

where $\forall n^v \to n^s$ is the mapping of all virtual node n^v onto substrate node n^s and $\forall l^v \to l^s$ is the mapping of all virtual link l^v onto substrate link l^s.

The virtual network embedding problem is divided into two sub-problems: virtual node mapping (VNoM) and virtual link mapping - VLiM. Firstly, in VNoM, a virtual network request i^{th} VNR_i is composed of a set of virtual segments $S_i^j \in VNR_i$ which normally include a coupe of virtual nodes and a virtual link between them. The node mapping N_i^v onto N^s is defined in Eq. 3:

$$M_{N_i^v} = N_i^v \to N^s \quad | \quad \forall n^v \in N_i^v \tag{3}$$

Constraint:

$$C(n^s) \geq \sum_{\forall n^v \to n^s} C(n^v) \tag{4}$$

Then, for each segment S_i^j with two virtual nodes n_a^v and $n_b^v \in S_i^j$, we have: $M_{n_a^v} = n_a^v \to n_a^s$ and $M_{n_b^v} = n_b^v \to n_b^s$. The constraint (5) ensures that all virtual nodes of a segment have to be located onto different substrate nodes.

$$n_a^s \neq n_b^s \mid n_a^v \to n_a^s, n_b^v \to n_b^s \tag{5}$$

Secondly, in VLiM, a virtual link l^v is mapped onto one substrate path (single path) or a set of substrate paths (multiple path) between two substrate nodes corresponding to the locations of mapped virtual nodes. The link mapping $M_{L_i^v}$ is defined as follows:

$$M_{L_i^v} = L_i^v \to P^s \quad | \quad \forall l^v \in L_i^v \tag{6}$$

Constraints of VLiM between virtual nodes n_s^v and n_d^v.

$$M_L^v(l^v(n_s^v, n_d^v)) \in P^s_{M_{N^v}(n_s^v), M_{N^v}(n_d^v)} \tag{7}$$

$$aB(P^s) \geq B(l^v(n_s^v, n_d^v)) \mid l^v(n_s^v, n_d^v) \rightarrow P^s \tag{8}$$

Constraint (7) ensures that the virtual link between n_s^v and n_d^v is mapped onto the substrate path that belongs to a set of available paths between two physical nodes that host these virtual nodes. And constraint (8) ensures that the available physical bandwidth between two physical hosting nodes is greater of equal the bandwidth request of virtual link.

Load. Each virtual segment in a virtual network request has the corresponding load for the physical infrastructure. An expression for the load of VNR_i is given by:

$$load_{VNR_i} = \sum_{\forall S_i^j \in VNR_i} \frac{\alpha \dfrac{C(n_x^v) + C(n_y^v)}{\sum_{\forall n^s \in N^s} C(n^s)} + \beta \dfrac{B(l^v)}{\sum_{\forall l^s \in L^s} B(l^s)}}{\alpha + \beta} \tag{9}$$

where:

- α, β are the factors for requirement of the CPU capacity of virtual nodes and bandwidth of virtual links, respectively. The default values $\alpha = \beta = 1$
- $C(n_x^v), C(n_y^v)$ are CPU requirement of virtual nodes n_x^v and n_y^v in segment S_i^j.
- $B(l^v)$ is the bandwidth requirement of virtual link l^v between n_x^v and n_y^v
- $\sum_{\forall n^s \in N^s} C(n^s)$ expresses total physical resource of substrate nodes.
- $\sum_{\forall l^s \in L^s} B(l^s)$ expresses total physical resource of substrate links.

3.2 Energy Profile of Network Device

A binary indicator $stt^n(t)$ represents the working state of substrate node with two values: 0 means this node is turned off and 1 otherwise. The energy consumption of the substrate network at time t is defined as $P_{net}(t)$ which represents the summing of all nodes with static (baseline) power, P_{st}, and power consumption of n_i physical links with ports, P_i, under their working speeds (Eq. 10).

$$P_{net}(t) = \sum_{\forall n \in N^p} stt^n(t).[P_{st} + \sum_{i=1}^{k} (n_i.P_i)] \tag{10}$$

For power measurement, we use energy profile of an energy-aware commercial HP 24×1 Gbps switch [10] which is able to change the working state of its port under separate operating speeds 1000 Mbps, 100 Mbps, 10 Mbps (Table 1).

Table 1. Power summary for a HP 24×1 Gbps switch

Operating speed	Power (W)
P_{st}	39
P_{10} - 100 Mbps per port	0.42
P_{100} - 1000 Mbps per port	0.48
P_{1000} - 1 Gbps per port	0.9

3.3 Objective

The main objective of this article is increasing energy-saving level of the network. The energy-efficient VNE algorithm focuses on reducing the number of working nodes by turning off as much as possible unused device. This objective is defined as following equation.

$$\min \sum_{\forall n \in N^p} state(n,t).[P_{static} + \sum_{j=1}^{k}(m_j.P_j)] \tag{11}$$

4 VNE Algorithms

Our strategy in this paper is *two state* embedding which deploys VNoM first and then, based on VNoM results, creates virtual links interconnecting mapped nodes on top of the substrate network. We propose two energy-efficient VNoM algorithms, namely *Energy-Efficient Capacity Greedy (EECG)* and *Reducing Middle Nodes (RMN)* algorithms. Thereafter, in VLiM, we use energy-efficient multipaths mappings combining with power scaling approach to provide suitable working speeds.

4.1 Energy-Efficient Capacity Greedy - EECG

Since the ILP approach for Virtual Network Embedding is NP-hard, we deploy the energy-aware node mapping based on Greedy algorithm. First of all, EECG sorts the substrate nodes in non-increasing order of available capacity, sort($n^s \in N^s$) by $aC(n^s)$. Then EECG chooses the node with the largest CPU capacity (CAP) value for mapping selection and turns all unused nodes after the mapping for energy-saving. The algorithm is described in more details in Algorithm 1. The $N_{nei}(n^V) \cap N_{vir}(n^S) \neq \emptyset$ constraint (line 10) ensures that the separate neighboring virtual nodes cannot be located onto the same substrate node. $N_{nei}(n)$ is a set of all neighbor of node n, where $N_{vir}(n^S)$ is a set of all virtual node that mapped onto the substrate node n^S.

Algorithm 1. Capacity Greedy node mapping

Input: N^S, N^V
Begin:
if $N^S = \emptyset$ *or* $N^V = \emptyset$ **then**
    MAPPING FINISH
else
    sort N^S by $aC(N^S)$ in non-increasing
    sort N^V by $C(N^V)$ in non-increasing
    for all $n^V \in N^V$ **do**
        for all $n^S \in N^S$ **do**
            if $N_{nei}(n^V) \cap N_{vir}(n^S)$ **then**
                continue
            else
                if $C(n^V) \geq aC(n^S)$ **then**
                    MAPPING FAILSE
                    break
                else
                    $N_{vir}(n^S) = N_{vir}(n^S) + n^V$
                    $aC(n^S) = aC(n^S) - C(n^V)$
                    MAPPING SUCCESS
MAPPING FINISH
Output: $N^S, N_{vir}(N^S)$

Algorithm 2. Node Ranking - *NRank*

procedure: NRank(N)
repeat:
Sort by $\mid N_{nei}(n) \mid$ in non-increasing, $\forall$ n $\in$ N
if $\mid N_{nei}(m) \mid = \mid N_{nei}(n) \mid \forall m, n \in N$ **then**
    Sort by $aC(n)$ in non-increasing
    if $aC(m) = aC(n), \forall m, n \in N$ **then**
        Sort by $\sum_{l \in L} B(l)$ in non-increasing
until N in non-increasing
end procedure

4.2 Reducing Middle Nodes Mapping

In NV, one virtual node can be mapped to only one substrate node, while a virtual link can be represented by a path which stays on group of consecutive physical links in substrate network. We consider that a *middle node* is an intermediate node in the substrate path mapped to a virtual link. Although the middle nodes are transparent to customers who request the VNRs, they are constantly consume energy. So that if we can reduce the number of middle nodes while satisfying the VNRs of customers, we can save the energy consumption of the network. RMN is a heuristic-based algorithm that focuses on minimizing the number of active nodes. The unused nodes are turned-off for reducing the energy consumption. To achieve this target, the mapping algorithm should minimize the number of active nodes at the VNoM step as well as avoiding purely forwarding

Algorithm 3. RMN node mapping

1: **Input**: N^s, N^v and *set of VNR*
2: **Begin**:
3: **if** $N^S = \emptyset$ *or* $N^V = \emptyset$ **then MAPPING FINISH**
4: **else**
5: Rank *VNRs* according to 1^{st} NRank$(N_i^v) \mid N_i^v \in VNR_i$
6: $NRank(N^S)$
7: Turn on $n^S \leftarrow pop(N^S), update(N_{on}^S, N_{off}^S)$
8: **for all** $VNR_i \in$ VNR **do**
9: $NRank(N_i^V) \mid \forall n^V \in N_i^V, N_i^V \in VNR_i$
10: **for all** $n^V \in N_i^V$ **do**
11: **for all** $n^S \in N^S$ **do**
12: **if** $N_{nei}(n^V) \cap N_{vir}(n^S) \neq \emptyset$ **then continue**
13: **else**
14: **if** n^S *is turn off* **and** $N_{nei}(n^S) \cap N_{on}^S = \emptyset$ **then continue**
15: **else** Turn on n^S $\&update(N_{on}^S, N_{off}^S)$
16: **if** $C(n^V) \leq aC(n^S)$ **then**
17: $N_{vir}(n^S) + n^V$
18: $aC(n^S) = aC(n^S) - C(n^V)$
19: **MAPPING SUCCESS**
20: $NRank(N_{on}^S)$
21: $N^S = N_{on}^S + N_{off}^S$
22: **MAPPING FINISH**
23: **Output**: $N^S, N_{vir}(N^S)$

middle nodes. In order to do that, we try to map a virtual link directly onto the substrate link (physical link that is directly connected between two substrate nodes). Intuitively, the number of substrate neighbors for node mapping partially guarantees the direct link mapping. Therefore we define the notion of node rank to evaluate the resource of nodes (*NRank* - Algorithm 2). The rank of a given node is determined by a number of its directly neighbor nodes, its CPU capability and collective bandwidth of its interfaces.

The *RMN* embedding algorithm (Algorithm 3) gets N^s and N^v as the inputs of the algorithms. The algorithm is stopped when either N^s or N^v becomes NULL. First of all, sets of substrate nodes and virtual nodes are sorted by *NRank*. We create two state-based subsets of substrate nodes, namely inactive nodes N_{off}^s and active nodes N_{on}^s.

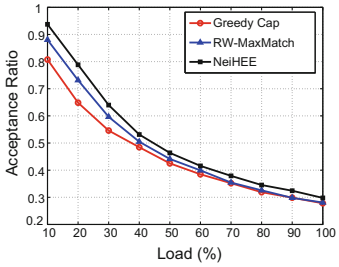
To start the algorithm, the first substrate node (highest-rank node) n^s in N^s is put into the active state and moved from a set of inactive nodes N_{off}^s to set of active nodes N_{on}^s. In each $VNR_i \in VNR$, we rank all nV and firstly try to find the suitable node in a set of active nodes N_{on}^S, then in a set of inactive nodes N_{off}^S. In case of turning on a node, it chooses the node that already has the active-state neighbors (line 15). Lines 26–27 describe the re-ranking process. With this algorithm, there are more opportunities to map a virtual link on a direct substrate link; also, middle nodes can be avoided.

4.3 Related Algorithms

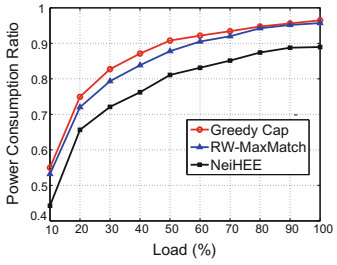
Topology-Aware: RW-MaxMatch. In [7] Xiang et al. proposed *RW-MaxMatch* that improves the acceptance ratio by using the node rank computing method. This ranking process is given by CPU power and collective bandwidth of a node. In order to estimate the acceptance ratio of our algorithm, we re-implement *RW-MaxMatch* and do comparison among these algorithms.

4.4 Performance Evaluation

As can be seen in the Figs. 1a and b, thanks to the *NRank* ranking that focuses on the number of neighbors, *NeiHEE* is not only better in terms of energy-saving level, but also maintains a good acceptance ratio in comparison with other mapping algorithms. In Fig. 1b the power consumption ratio is one when all switches and links are turning on in the maximum working rate.



(a) Acceptance rate



(b) Power consumption ratio

Fig. 1. Acceptance ratio and power consumption among three algorithms

5 Conclusion and Future Work

In this paper we propose two energy efficient virtual network embedding algorithms, Greedy Capacity and Reducing Middle Node. The algorithm are implemented and evaluated in term of power saving level as well as investigate their

performance by increasing acceptance ratio. As we can see in the experimental results, the RMN algorithm can reduce the power consumption while Greedy Capacity embedding algorithm is lightweight. For the future work we intend to extend our algorithm in the real testbed and evaluate the QoS values. The result will help us to efficiently trade-off between power consumption, quality and performance of the system.

References

1. Global e-Sustainibility Initiative (GeSI) – SMART 2020: enabling the low carbon economy in the information age (2012)
2. Sherwood, R., Gibb, G., Yap, K., Appenzeller, G., McKeown, N., Parulkar, G.: "FlowVisor: a network virtualization layer" OpenFlow Switch Consortium, Technical report OPENFLOW-TR-2009-1 (2009). http://www.OpenFlow.org/wk/index.php/
3. Ngoc, N.P., Huu, T.N., Quang, T.V., Hoang, V.T., Thu, H.T., Tran-Gia, P., Schwartz, C.: A new power profiling method and power scaling mechanism for energy-aware NetFPGA gigabit router. Comput. Netw. **78**, 4–25 (2015)
4. Botero, J.F., Hesselbach, X., Duelli, M., Schlosser, D., Fischer, A., de Meer, H.: Energy efficient virtual network embedding. IEEE Commun. Lett. **16**(5), 756–759 (2012)
5. Houidi, I., Louati, W., Ameur, W.B., Zeghlache, D.: Virtual network provisioning across multiple substrate networks. Comput. Netw. **55**, 1011–1023 (2011). Telecom SudParis, 9 rue Charles Fourier, 91011 Evry, France
6. Botero, J.F., Hesselbach, X.: Greener networking in a network virtualization environment. Comput. Netw. **57**, 2021–2039 (2013). Universitat Politcnica de Catalunya, Jordi Girona Street, 1 and 3, Barcelona, Spain
7. Cheng, X., Sen, S., Zhang, Z., Wang, H., Yang, F., Luo, Y., Wang, J.: Virtual network embedding through topology-aware node ranking. Comput. Netw. **56**(6), 1797–1813 (2012). doi:10.1016/j.comnet.2012.01.022
8. Nguyen, H.T., Vu, A.V., Nguyen, D.L., Nguyen, V.H., Tran, M.N., Ngo, Q.T., Magedanz, T.: A generalized resource allocation framework in support of multi-layer virtual network embedding based on SDN. Comput. Netw. **92**, 251–269 (2015). http://doi.org/10.1016/j.comnet.2015.09.042
9. Bolla, R., Bruschi, R., Davoli, F., Cucchietti, F.: Energy efficiency in the future internet: a survey of existing approaches and trends in energy-aware fixed network infrastructures. IEEE Commun. Surv. Tutorials **13**(2), 223–224 (2011). doi:10.1109/SURV.2011.071410.00073
10. Mahadevan, P., Sharma, P., Banerjee, S., Ranganathan, P.: Energy aware network operations. In: Proceedings of INFOCOM (2009)

Research on Enhancing Security in Cloud Data Storage

Minh Le Quang[1(✉)], Phan Huy Anh[1], Nguyen Anh Chuyen[2], and Le Khanh Duong[2]

[1] Information Technology Institute, Vietnam National University, Hanoi, Vietnam
lqminh78@gmail.com, phananh21588@gmail.com
[2] University of Information and Communication Technology, Thai Nguyen University, Thai Nguyen, Vietnam
nachuyen@ictu.edu.vn, duonglk@gmail.com

Abstract. Nowadays, cloud storage service has been widely deployed by multiple providers that offer free and large-sized data storage. Nevertheless, along with the advantages are the risks of data loss, the service providers' unauthorized access to data, and data loss due to the providers' unsustainability. This paper will investigate how to use cloud storage capability in a safe and low-cost way. The authors will present solutions for secured cloud data storage based on the RAID redundancy mechanism, partially encrypted data and mathematical models established on the probability theory and the system reliability.

Keywords: Cloud storage · RAID · Mathematical models · Reliability

1 Introduction

Cloud computing in general or cloud storage in particular is a revolution in information technology. According to IEEE Computer Society, "Cloud Computing is a paradigm in which information is permanently stored in servers on the internet and cached temporarily on clients that include desktops, entertainment centers, table computers, notebooks, wall computers, handhelds, sensors, monitors, etc." Cloud storage has become a service, as a means to replace traditional storage from like a USB, a memory card, or even a hard disk drive. This is the most optimal data storage form because of its flexibility: Most of them support multi-platform (Windows-based, Mac, Linux, iOS, Android, etc.) Besides its advantages, cloud storage has its own drawbacks, for example risks and information security [4,5].

Nowadays, there are a number of cloud storage providers which offer great service to users like Dropbox, SkyDrive, Box, Google Drive, and Amazon Cloud Drive, etc. Users can choose either free or subscribed service with storage capacity ranging from 2G to 50 GB or unlimited. Data can be synchronized across several platforms including Web, Mobile and Desktop. However, it is critical to

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_55

concern about data security since there are threats such as account loss, lifespan of service providers, and losing access to accounts due to providers' system being hacked. In this paper, authors will propose a method to enhance data security in cloud storage, provide mathematical models stabled on the probability theory.

2 RAID Mechanism and Data Security in Cloud Storage

2.1 RAID Features in Data Storage

RAID Definition. RAID (Redundant Arrays of Inexpensive Disks or Redundant Arrays of Independent Disks) is a technology that combines independent physical disk drives into a single hard drive for the purpose of read/write speed improvement or reliability of stored data enhancement, or both [3].

There are 2 implementations of RAID:

- Hardware RAID requires a RAID controller that controls Input/Output. Hardware RAID is used for host servers. It is high-performance yet expensive.
- Software RAID: The operating system controls Input/Output. Software RAID is implemented on computers to boost the performance with low-cost solution.

Three reasons to implement RAID are: backup, high performance, low cost. RAID is a critical component to host servers in data storage.

RAID Levels. RAID Advisory Board (RAB), founded in July 1992 with the purpose of RAID education and standardization has classified RAID levels and hardware standard. According to RAB, there are 7 levels of RAID with different features, established on two basis levels RAID 0 and RAID 1.

(a) RAID 0

RAID 0 consists of at least 2 similar disks, which creates an array of n disks ($n \geq 2$). Data is split up evenly and get written across all devices in the array. Each disk stores 1/n data. The size of the array is the size of the smallest drive multiples the number of drives [3].
Array Capacity = Size of Smallest Drive * Number of Drives.
The capacity of the array is the sum of two drives' capacity.
For example: If we use two 80 GB drives, we have an 160 GB array.
Advantages: Read/Write transfer rate enhancement: Each disk has to Read/Write 1/n of the data. Theoretically, performance is n times higher.
Disadvantages: Lower reliability. If one drive fails, all data in RAID 0 array is lost. Data loss rate of RAID 0 array is n times higher than the single-disk's (Fig. 1).

(b) RAID 1

This is the simplest RAID level that provides data reliability. Like RAID 0, RAID 1 requires at least 2 drives to operate. Data are stored twice in 2 drives

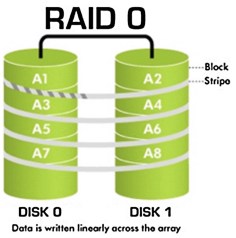


Fig. 1. RAID 0 data storage

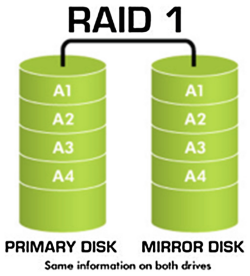


Fig. 2. RAID 1 data storage

(Mirroring). If one disk fails, the other will continue to operate. Therefore, broken drive can be replaced without any worry of data loss [3].

RAID 1 is not high-performance; however, it is essential for administrations and individuals that manage important data.

An RAID 1 array capacity is the size of a single drive. (For example: If an RAID 1 array consists of 2 80 GB drives, the array will be 80 GB in capacity) (Fig. 2).

(c) RAID 10

RAID 10 combines the approaches of RAID 1 and RAID 0. It requires a minimum of 4 drives to set up an RAID 10 array. Data are written on 4 drives at the same time: using Striping (RAID 0) on 2 drives, and using Mirroring (Raid 1) on the two others.

RAID 10 is fast and secure. Performance are improved while reliability is ensured even if 1 drive fails.

However, RAID 10 has its disadvantages of high cost, effective space is 1/2 of total size of 4 drives (which is the same as RAID 1's drawback) [6] (Fig. 3).

Besides the above mentioned levels, there are RAID 2, RAID 3, RAID 4, RAID 5 and RAID 6. RAID supports each implementation with various features. Although each is different, they are all advanced from the traditional RAID.

2.2 Security of Cloud Storage Issues

Data security and preservation in cloud storage is critical to businesses – users. Securing cloud computing consists of 2 aspects: security methods for service

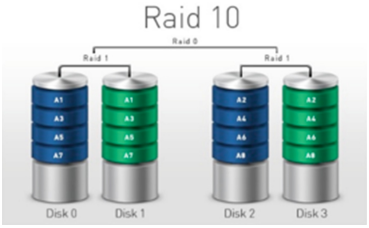


Fig. 3. RAID 10 data storage

providers and for users [1].

Cloud computing storage is currently the attractive target to cyber-criminals (according to European Union Agency for Network and Information Security - ENISA) [2]. ENISA reports top 10 threats to cloud computing security, including:

- Code injection attacks (e.g.: SQLi and Directory Traversal)
- Data stealing/collecting Worms and Trojans
- Drive-by exploits (utilizing vulnerabilities in web browsers to attack)
- Information leakage exploits, especially with the outburst of mobile device usage
- Confidential information intrusion from internal or external
- Attacks through remotely controlled computer networks (Botnets)
- Denial of service (DDoS)
- Identity theft
- Human threats (knowledge shortage or staff's discontent)
- Targeted attacks

Data synchronization across all user's devices is wonderful features of cloud storage service. Nonetheless, with its convenience, comes the potential threats to users. If hackers can control one employee's account, they can easily distribute infected document to the whole organization with synchronization feature. According to Kaspersky Lab's analysis, 30 % of malicious software explored in cloud folder on user's personal computer at home are infected through synchronization. It is 50 % with business users.

Hackers can access and steal information without any realization from customers by exploiting vulnerabilities of service providers. This type of attacks is named "man-in-the cloud" – the common vulnerabilities of almost all cloud storage service [7].

Users' data has also become government's target. In 2012, Google received more than 21.000 requests from U.S government to provide information of more than 33.000 users. Other technology businesses like Microsoft also received more than 70.000 requests on 122.000 user accounts in the company storage system. Those numbers are extracted from Lucas Mearian's analysis on cloud storage security [8].

Data loss threats are not solely providers' problem but also customers'. Users are hackers' prey to exploit information by impersonating. They exploited vulnerability of Google Drive - a Google's reliable cloud computing service - to collect victim's data. By using this method, a Middle East hackers group deployed a wide-scale cyber-fraud on July 2015 [9].

Data insecurity, security breaches, service provider's security methods and users' neglect are utilized by hackers. There are always various threats to data. Originated from data storage security when using storage services and RAID models of data storage, the authors would like to propose a secure data cloud storage established on RAID mechanism, partially encrypted data and mathematical models based on the probability theory and the system reliability.

3 Solution for Enhancing Data Security in Cloud Storage

3.1 RAID Enhanced Security Cloud Storage - RESCS Proposal

RESCS Proposal. RESCS – RAID Enhance Security Cloud Storage is data storage mechanism on cloud services that authors would like to introduce. RESCS uses free service from providers like Google Drive, Dropbox, Box, and OneDrive, etc. RESCS combines the secured storage with backup of RAID 0, 1, the flexibility of cloud storage service, along with partial data encryption, and probability-theory-based and system-reliability-based mathematical models to provide solution that can solve 2 major problems of data cloud storage:

- **Integrity:** Data are stored on a number of service provider servers, and not completely dependent on any providers; therefore, when one accounts is inaccessible, others can continue to operate. Then lost or corrupted data will be restored using the solution distribution mechanism.
- **Security:** Using data fragmentation to split up data into different blocks to store on different providers' cloud, RESCS ensures the safety of information from attacks or unauthorized access from service providers. Even when all components of the data are gathered by users, data are inaccessible because each component is encrypted.

RESCS Data Storage Mechanism. RESCS uses free services from cloud service providers like Google drive, Dropbox, OneDrive, and Box, etc. to store data. Free accounts are created with user's email. To assure the integrity of data, it requires user to register to at least 3 providers service and a minimum of n accounts in each cloud ($n \geq 2$). In short, the numbers of accounts for storage is 3*n (Fig. 4).

Data will be stored in cloud services by following process: RESCS splits up user's data and partially encrypts them then stores data into different accounts with the same method as RAID 10 model. For example: data are split up into 9 blocks and stored in 3 accounts on each cloud service (Fig. 5).

User's data are stored on cloud accounts based on the following rules:

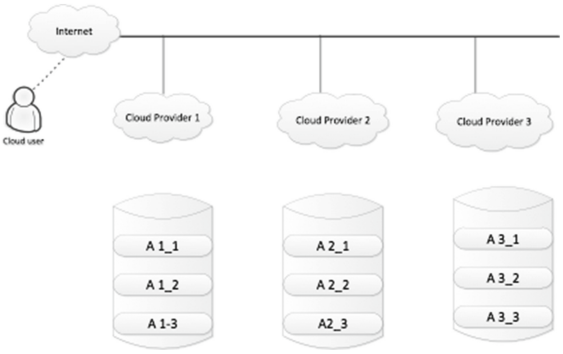


Fig. 4. RESCS storage mechanism

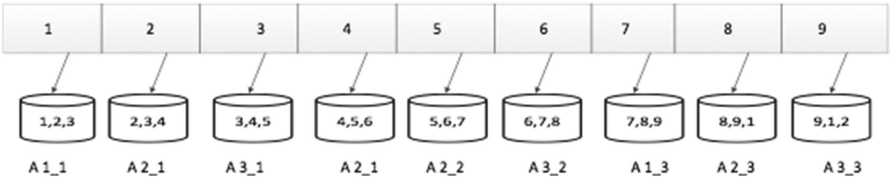


Fig. 5. Data fragmentation and storing on cloud services

- Accounts from 1 provider are placed alternately as the model: n*i+m (n: the number of accounts in on 1 service, i: the number of turn, m: the order of the account)
- On each account, three data blocks are stored adjacently based on the fragmentation and encryption order.
- The first and the last block are stored on the same account.

The above rules for fragmentation and storing have below listed advantages:

- When one account is lost or inaccessible, data can be retrieved from any 2 adjacent accounts.
- When data on one service is lost, user can use backup data from accounts on other services. If both providers fail to operate, data will not be restored.

The next aspect of the solution is how to manage the order of accounts used for storage and the order of data blocks. The order of accounts can be variable to complicate the hacking attempt. Most of service providers limit the storage capacity. It varies across providers: 2 GB on Dropbox, 15 GB on Google Drive, 5 GB on Box and OneDrive, etc. The maximum size of a file uploaded to a service is also limited. However, it depends on transfer speed, technological infrastructure, and data security, etc. Therefore, with RESCS, authors recommend the maximum file size is 200 MB.

Due to the differing file sizes, to ensure the security for data on cloud services, RESCS fragments data based on the number of accounts or the number of cloud

services to optimize storage of small file. After fragmenting data, RESCS will include the header with the following structure in each file for management purpose (Fig. 6):

Total package	Order Package	Next Storage	Filesize	Data….

Fig. 6. Data block header structure

In which:

- Total package: the total number of blocks that the data are split into
- Package order: the order of the block in the structure
- Next storage: the place where the next block is stored
- File size: the size of the block, used to check when conjoining all blocks
- Data: the data of the block

As data are fragmented and dispersed on independent services, each block when accessed will not reveal the integrated data. Nevertheless, fragmented data blocks of simple files without header structure like txt can still be exploited. Hence, data encryption is vital.

Mathematical Model Based on Probability Theory and Reliability of the System. Input data is one user's file stored in cloud storage service with a particular size (Fig. 7).

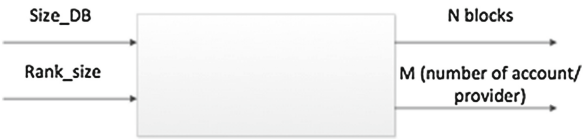


Fig. 7. RESCS operation model

Size_DB: the size of input data
Rank_size: the maximum size of uploaded file
N: the number of fragmented blocks
M: the number of account on cloud storage service

Input data file size is Size_DB. Since RESCS uses different accounts on independent services and each provider has its own policy on file size limit, the maximum file size for each data block is the smallest file size limit (Rank_size) among all service providers. Data are split into N blocks and distributed evenly for the number of account on each services (M). Suppose that N = M.
P is initial reliability of the system. (P_{ss}).

Suppose that data are split into 9 blocks, 3 providers are chosen with 3 accounts on each, placed alternately. α_1, α_2, α_3, β_1, β_2, β_3, γ_1, γ_2, γ_3 the reliability of each block. Then the initial reliability of the system is:

$$(P_{ss}) = \alpha_1 * \alpha_2 * \alpha_3 * \beta_1 * \beta_2 * \beta_3 * \gamma_1 * \gamma_2 * \gamma_3 \tag{1}$$

Suppose that: $\alpha_1 = \alpha_2 = \alpha_3$
$\beta_1 = \beta_2 = \beta_3$
$\gamma_1 = \gamma_2 = \gamma_3$

Then the initial reliability of the system is:

$$(P_{ss}) = \alpha^3 * \beta^3 * \gamma^3 \tag{2}$$

Case 1: Each account stores 2 data blocks.

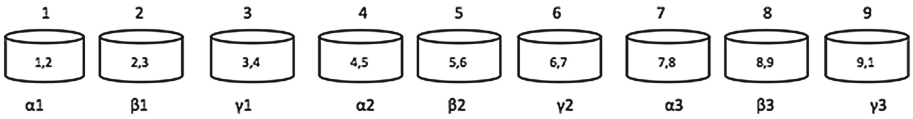


Fig. 8. The reliability of system in Case 1

In this case, when interleaved accounts are loss, data can be restored from the surrounding accounts. If provider stops the service, data will be safe thanks to adjacent accounts (Fig. 8).
The initial reliability of the system in case 1 is (P_{ss1}), then:

$$(P_{ss1}) = \alpha^3 * \beta^3 * \gamma^3 + (1-\alpha)^3 * \beta^3 * \gamma^3 + \alpha^3 * (1-\beta)^3 * \gamma^3 + \alpha^3 * \beta^3 * (1-\gamma)^3 + \alpha * \beta^2 * \gamma^2 * (1-\alpha)^2 * (1-\beta) * (1-\gamma) + \alpha^2 * \beta^2 * \gamma * (1-\alpha) * (1-\beta) * (1-\gamma)^2 + \alpha^2 * \beta * \gamma^2 * (1-\alpha) * (1-\beta)^2 * (1-\gamma) \tag{3}$$

Case 2: Each account stores 3 data blocks.

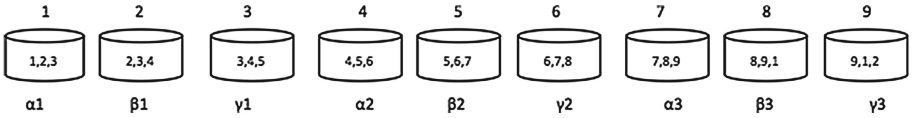


Fig. 9. The reliability of system in Case 2

In Case 2, if 2 adjacent accounts are lost or inaccessible, data can be retrieved from neighboring accounts. If one provider stops service, data from neighboring

accounts can be used instead (Fig. 9). If (P_{ss2}) is the reliability in Case 2, then:

$$(P_{ss2}) = \alpha^3 * \beta^3 * \gamma^3 + (1-\alpha)^3 * \beta^3 * \gamma^3 + \alpha^3 * (1-\beta)^3 * \gamma^3 + \alpha^3 * \beta^3 * (1-\gamma)^3 + (1-\alpha)^3 * (1-\beta)^3 * \gamma^3 + \alpha^3 * (1-\beta)^3 * (1-\gamma)^3 + (1-\alpha)^3 * \beta^3 * (1-\gamma)^3 \quad (4)$$

Suppose that this two cases have high reliability, but if one service changes its security policy or at that time hackers are exploiting vulnerability, reliability of the system will be significantly reduced. The redundancy of this two cases will become more feasible by improving the reliability to 1.56 % and 3.7 %.

Example: With $\alpha = 0.9999$ rate of error account is $^1/_{10000}$, $\beta = 0.9999$ rate of error account is $^1/_{10000}$, $\gamma = 0.8$ rate of error account is $^{2000}/_{10000}$, the reliability of both 2 cases is improved by 1.56 %. With $\alpha = 0.99999$ rate of error account is $^1/_{100000}$, $\beta = 0.99999$ rate of error account is $^1/_{100000}$, $\gamma = 0.75$ rate of error account is $^{25000}/_{100000}$, the reliability of both 2 cases is improved by 3.7 %. When using these models, the reliability is improved, so data integrity is warranted.

3.2 RESCS Evaluation and Comparison

The solution ensures the security of user's data even when their password is stolen or violated by providers. Data encryption is the key for the issue. Nowadays, there are applications that encrypt user's data before uploading it to cloud storage.

Credeoncp is a client-side encryption application for cloud storage. It is compatible with all major cloud service providers. The application encrypts user's files to protect them from unauthorized access even from the government. It offers AES 256 and FIPS 140-2 encryption.

Spideroak is an application that allows users to store data on the cloud. Files will be encrypted by user's password before being uploaded to the server. Password information is then saved on user's computer at home. Nothing is stored on the provider's server. This ensures confidentiality of user data because even service providers will not be able to access data without a password.

Box Cryptor works as an agent between user and cloud storage providers like Dropbox, Google Drive, and OneDrive, etc. The application encrypts user's data before uploading them onto cloud services. Data is accessible on different platforms like mobile and desktop and operating systems like Windows, MAC, and Linux.

These above mentioned solutions ensure data security yet not data integrity. RESCS offers both by using partial data encryption.

4 Conclusion and Directions for Future Research

Cloud computing technology is growing rapidly, becoming a future trend and a new revolution in Information technology. This platform is widely used for complex computing applications and data cluster formation. Data security has been

a concern and subject of many researches. In this paper, authors introduce arguments and evidences on data insecurities from both users and service providers of cloud data storage. From those, authors propose solution for secure data storage based on RAID mechanism. Although the solution is solely theoretically presented in the paper, it resolves the major problems of cloud storage service which are data security and integrity. The paper, at the same time, provides mathematical models based on probability theory to calculate the availability and fault tolerance of the system.

For future research, the group of authors will focus on modeling the solution and systematizing RESCS for installation testing. Alongside, authors will conduct research on an encryption mechanism for fragmented data before being uploaded to the cloud server, as well as synchronizing mechanism across devices and operating systems like several storage services are providing.

Acknowledgements. We would like to express our thankful to the support of project QCT.16.01 of Information Technology Institute - Vietnam National University, Hanoi.

References

1. Catteddu, D., Hogben, G., Benefits, C.: Risks and recommendations for information security. In: European Network and Information Security Agency (ENISA) (2009)
2. Chhibber, A., Batra, S.S.: Security Analysis of Cloud Computing. Innovation Labs, Tata Consultancy Services Ltd., Kolkata, India (2013)
3. Glenn Berry, C.: SQL Server Hardware. Simple Talk Publishing (2011). ISBN: 978-1-906434-62-5
4. Subra Kumaraswamy, M.S.: Introduction to Cloud Security Architecture from a Cloud Consumer's Perspective (2011)
5. M.S.: ENISA:Threat Landscape: Responding to the Evolving Threat Environment (2013)
6. RAID Levels and SQL Server. https://technet.microsoft.com/en-us/library/ms190764(v=sql.105).aspx
7. Whittaker, Z.: Attackers can access Dropbox, Google Drive, OneDrive files without a user's password (2015). http://www.zdnet.com/article/dropbox-google-drive-onedrive-files-man-cloud-attack
8. Reilly, C.: Hackers hold 7 million Dropbox passwords ransom (2014). http://www.cnet.com/news/hackers-hold-7-million-dropbox-passwords-ransom

Resource-Aware Scheduling in Heterogeneous, Multi-core Clusters for Energy Efficiency

Xuan T. Tran[1,2(✉)]

[1] Faculty of Electronics and Communication Technology,
Information and Communication Technology University, Thai Nguyen, Vietnam
ttxuan@ictu.edu.vn

[2] Analysis, Design and Development of ICT Systems (AddICT) Laboratory,
Budapest University of Technology and Economics, Budapest, Hungary

Abstract. The benefits and necessity of multi-core technology are undeniable and make it a critical trend in chip manufacture. This shift, however, also brings complexities in computer sciences, especially in job scheduling problem. Additionally, energy bills have been a major concern due to the increasing population of computing systems lately. The trade-off between performance and energy efficiency in such systems makes the scheduling optimization more challenging.

This study aims to propose an energy-efficient scheduling solution that exploits the resource heterogeneity and utilization in computing clusters of multi-core processors. The numerical results show that the proposed policy helps saving significant energy in a heterogeneous cluster.

Keywords: Multi-core · Heterogenous resources · Energy efficiency · Resource-aware scheduling

1 Introduction

Multi-core and multi-processing technologies have been first introduced in special and embedded systems for a long time. With its tremendous benefits [1,2,4], multi-core architecture has become the *de factor* design for commercial chips for PCs, mobile phones, servers, etc. These technologies, however, increase the heterogeneity and complexity of computing systems and bring new challenges faced by service providers. Among challenges, job scheduling is an NP-hard optimization problem in such systems due to several criteria [5–8].

In addition, green computing lately became a major concern due to the quick increase in sizes of modern systems. The availability of Dynamic power management (DPM) techniques such as switching off/on components and dynamic voltage and frequency scaling (DVFS) makes it fascinating for power savings [9–13]. However, DPM techniques come with the trade-off between performance and energy efficiency. This work aims to propose a solution for energy-efficient scheduling in multi-core computational systems. The job scheduler is aware of resource heterogeneity and utilization to make decision about job assignment.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_56

A simple, best-effort hottest resource allocation policy (last-in, first out) is applied by the scheduler for comparison.

The rest of paper is organized as follows. Section 2 presents a brief description for related works and the approach of this study. Numerical results and evaluations are given in Sect. 3. Lastly, Sect. 4 concludes the paper.

2 Case of Study

2.1 Related Works

Literature studies focused on different aspects related to job scheduling problem in distributed multi-core systems. Some previous works approached optimizing system performance with scheduling heuristics [7,14,15]. In [8], authors investigated the scheduling issue for atomic jobs in multi-core computational clusters and addressed that parallel processing is a resource-efficient technology in the such system. The study also took into account the power management technique at different levels for energy savings. Sergio et al. approached an energy-aware scheduling by providing new formulation of job makespan and energy consumption for a 2-level grid system [3]. Energy-efficient scheduling policy based on the heterogeneity of cluster resources was studied in [16,17] which inspired our approach.

2.2 System Model

A heterogeneous cluster is constructed with S types of commercial multi-core servers and exploits the common-queue buffering scheme [17], as illustrated in Fig. 1a. Computing servers have parameters that follow `SPECpower_ssj2008` benchmark of the Standard Performance Evaluation Corporation (SPEC) [18].

Considered parameters of a server with type s ($s \in S$) are:

- N_s (the number of homogeneous cores),
- C_s (the `ssj_ops` value, defined by the number of operations per second),
- $P_{ac,s}$ (the average active power),
- $P_{id,s}$ (power consumption when server is active idle), and
- $EE_s = C_s/P_{ac,s}$ (the Energy Efficiency (EE) ratio).

We consider gang workload (see Fig. 1b) with the following characteristics:

- each job is composed of a gang of tasks, which arrive and depart together. Each task requires one core for its execution;
- a job's tasks can be executed on any available cores of any server;
- jobs have service time demand unknown to the scheduler;
- jobs are independent and non-preemptible.

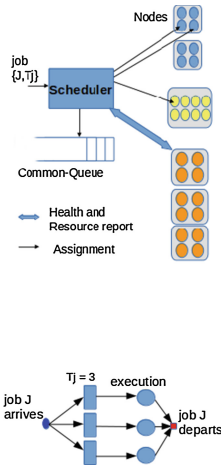


Fig. 1. (a) A model of heterogeneous, multi-core cluster (b) A gang job example:job J is split into $T_j = 3$ sibling tasks which must be processed simultaneously

2.3 Scheduling Policies

Let jobs attend the scheduler with the First Come First Served (FCFS) service policy and require a number of CPUs for execution. The scheduler keeps track about available resources in the cluster and is responsible to allocate an appropriate amount of resources to an incoming job.

The two policies considered in this study are described as the following:

- *LIFO policy*: The scheduler maintains the list of free servers as a stack and implements the last-in, first-out (LIFO) allocation. When a job requests resources for its tasks, if there exist enough computing cores for the job, the scheduler assigns tasks to the firstly found (on-top) node in the list. The process is iterated until all tasks are assigned. Pseudo-code implementing this policy is presented in Algorithm 13
- *Resource-aware policy*: The scheduler keeps track available resources. In contrast to LIFO policy, the scheduler maintains the list in the order of the energy efficiency (EE) values such that servers with higher EE values are prior. Another constraint used for the allocation is the resource utilization of servers in the list. The scheduler will priorly choose servers which are running and have higher energy efficiency values. This avoids to wastefully utilize inactive servers while running servers remain unused resources. This resource allocation policy is illustrated in Algorithm 14

Let a job be defined by $\{j, t_j\}$, where j is job identification and t_j is the number of tasks, which require an equal performance capacity of one CPU. That means job j needs t_j cores simultaneously for its execution. Tasks of the job can be distributed across the cluster and thus have different processing times.

Therefore, the other tasks still occupy the resources until the last task is finished and job service time equals to the execution time of the longest task.

Algorithm 13. LIFO algorithm

For an incoming job (j, t_j)
Input: L is the list of free servers with size $|L|$
Output: R is the list of allocated servers
Initiate: $R = \emptyset$, $allocated_cores = 0$
if sum of free cores in L is LESS than t_j or queue is not EMPTY **then**
    INSERT job (j, t_j) to waiting queue
else
    while $i = L - 1$ and $i \geq 0$ **do**
        MOVE $server[i]$ to the list R
        DECREMENT the length of L
        ADD $server[i].free_cores$ to $allocated_cores$
        if $allocated_cores \geq t_j$ **then**
            job $(j, t_j) \leftarrow R$

2.4 Measurement Metrics

Response time of system for jobs is one of the most typical metrics to measure the performance. Additionally, waiting time in queue of jobs indicates the actual delays for job processing caused by system. Let w_l be waiting time in queue of job l before it is executed and s_l be service time that takes the system to process it. The response time r_l of job l is the time period from its arrival till its departure instant, thus:

$$r_l = w_l + s_l \tag{B.1}$$

The average waiting time and response time per job after n jobs completed can be formulated as follows respectively:

$$WT(n) = \frac{\sum_{l=1}^{n} w_l}{n}, \tag{B.2}$$

$$RT(n) = \frac{\sum_{l=1}^{n} r_l}{n}, \tag{B.3}$$

Energy consumption of a server during the system operation is calculated as the sum of idle energy consumption (when a server stays idle) and active energy consumption (when the server is processing any tasks of jobs). As a server can partially run a subset of cores or all cores in a time period, the active power changes according to the actual number of running cores, as formulated

Algorithm 14. Resource-aware (RA) algorithm

For an incoming job (j, t_j)
Input: L is the list of free servers in the order of energy efficiency values with size $|L|$ (server $L-1$ has the highest energy efficiency)
Output: R is the list of allocated servers
Initiate: $R = \emptyset$, $allocated_cores = 0$
if sum of free cores in free server list is LESS than t_j or queue is not EMPTY **then**
    INSERT job (j, t_j) to waiting queue
else
    while $i = L - 1$ and $i \geq 0$ **do**
        if $server[i]$ is RUNNING **then**
            MOVE $server[i]$ to R
            ADD $server[i].free_cores$ to $allocated_cores$
            DECREMENT the size of L
            if $allocated_cores \geq t_j$ **then**
                job $(j, t_j) \leftarrow R$
    if $allocated_cores$ do not fulfill t_j **then**
        for i in all remaining servers of L **do**
            MOVE $server[i]$ to R
            ADD $server[i].free_cores$ to $allocated_cores$
            if $allocated_cores \geq t_j$ **then**
                job $(j, t_j) \leftarrow R$

in [8]. The measurement of active energy consumption takes account actual active power of a server during the runtime. Let $oe(n)$ be total energy consumed by system during the operation of n completed jobs. The average operating energy consumption per job is calculated as follows.

$$AOE(n) = \frac{oe(n)}{n}, \tag{B.4}$$

All considered parameters and metrics are listed in Table 1.

3 Simmulation and Numerical Results

The study approach is built in a simulation software using C program. The simulation is run with applied confident level of 99 % and the stop condition at of three million completions. The specifications of used servers and applied work loads are given in Subsect. 3.1 Obtained results are illustrated in Subsect. 3.2.

3.1 Server Specifications and System Loads

In run simulation, considered cluster is composed by $S = 4$ types of servers. Chosen server types and the server number for each type are presented in Table 2. It is noted that computing capacity of server type s, C_s presented for ssj_ops value, and its power consumption, $P_{ac,s}$, are measured at 100 % target load. As

Table 1. Notations

S	Number of server types
M_i	Number of servers of type i
λ	system arrival rate
μ	service rate of entire system
μ_i	service rate of a computing core in a server type i
ρ	Average system utilization
t_j	Number of task of job j
WT(n)	Average waiting time per job up to n job completed
RT(n)	Average response time per job up to n job completed
AOE(n)	the average energy consumption per job up to n job completed

Table 2. Server parameters

Server type	N_*	C_*	$P_{ac,*}$ (W)	$P_{id,*}$ (W)	EE_*	No. of servers
FUJITSU TX1320 M1 (Intel E3-1275L v3) [20]	4	508794	60	13.3	7535	10
Acer AR380 F2 (Intel E5-2640) [19]	12	990555	254	81.9	3904	10
Fujitsu TX1330 M2 (Intel E3-1240L v5) [22]	4	484122	47.2	16.1	8187	20
Acer AR380 F2 (Intel E5-2665) [21]	16	1432448	258	72.4	4909	20

chosen servers have homogeneous multi-cores, the performance capacity of one core in each server type is calculated as in Table 3.

Assuming that job's inter-arrival time and service time are exponentially distributed with mean $1/\lambda$ and $1/\mu$, respectively. We also suppose that the number of tasks composing a gang job is uniformly distributed in the closed interval of [1,20]. Thus the average number of tasks per job is $T_{avg} = (1 + 20)/2 = 10.5$. It means that each job requires an average resource amount of 10.5 cores for its execution. Let μ_1, μ_2, μ_3 and μ_4 denote the service rates of types Intel E3-1275L v3, Intel E5-2640, Intel E3-1240L v5, and Intel E5-2665 respectively. A job task is assumed requiring a computing capacity equivalent to 127198.5 ssj_ops in average (see Table 3). That means if a task is assigned to a core of server type Intel E3-1275L v3, Intel E5-2640, Intel E3-1240L v5, or Intel E5-2665, it will be served with service rate of $\mu 1 = 1/s$, $\mu_2 \approx 0.65$, $\mu_3 \approx 0.95$, or $\mu_4 \approx 0.70$ respectively.

Table 3. Capacity per core

Server type	$C_*/core$
Intel E3-1275L v3	127198.5
Intel E5-2640	82546.25
Intel E3-1240L v5	121030.5
Intel E5-2665	89528

Service rate of the whole system can be calculated as follows:

$$\mu = \frac{\sum_{i=1}^{S} M_i \times N_i \times \mu_i}{T_{avg}}, \tag{B.5}$$

and traffic volume carried by system is defined by the ratio between the arrival rate and service rate as:

$$\rho = \frac{\lambda}{\mu}. \tag{B.6}$$

We run with traffic volume values of 30 %, 40 %, 50 %, 60 %, and 70 %. Following Eqs. (B.5) and (B.6), we find system service rate is approximate to 39,8 jobs/s and arrival rates are approximate to: 11.94, 15.92, 19.90, 23.88, and 27.86, respectively.

3.2 Obtained Results

Figure 2 shows the average operating energy consumption of system for a job execution, regarding to system load. The proposed solution remarkably yields energy efficiency, compared to the LIFO policy. Particularly, it saves about 70 % of energy at low load and roughly 57 % at high load.

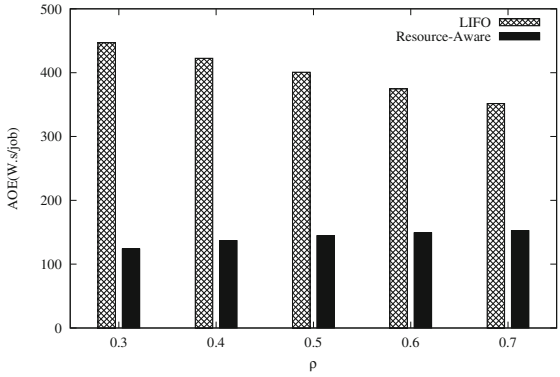


Fig. 2. Average operating energy consumption per job (KW.s/job)

System performance metrics, in terms of average waiting time and average response time per job, are shown in Figs. 3a and b. The average waiting time indicated that the delay in queue of jobs reduces when applying resource-aware policy. Figure 3b also shows a better service of the system with about 25 % response time reduction if the proposed solution is used instead of LIFO policy.



Fig. 3. (a) Average waiting time per job (s/job) (b) Average response time per job (s/job)

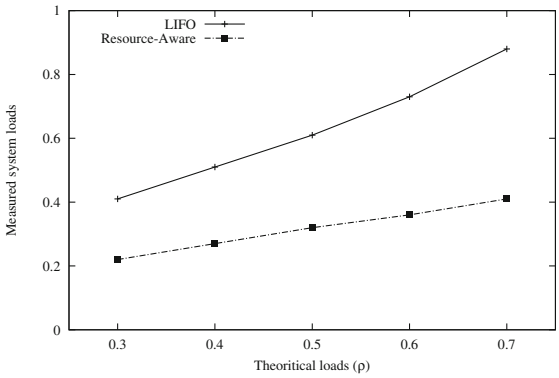


Fig. 4. Measured loads vs. theoretical loads

We plotted the measured system traffic volume in Fig. 4. It is shown that the proposal helps system utilize the resources more efficiently and thus keep the overall load lower compared to theoretically calculated load, whist LIFO policy yields an oposite result. We should note that the obtained results highly depend on the parameters of chosen servers. However, the proposed resource-aware policy is worth for the energy efficiency of any heterogeneous, multicore systems.

4 Conclusion

This study investigated an approach for energy-efficient scheduling policy that is based on the awareness of resource availability and heterogeneity. The proposal was compared to a traditional best-effort LIFO policy.

The numerical results indicate that the proposed policy yields significant energy savings in the system whilst the performance is not traded off. Furthermore, system outperforms when the resource-aware approach is applied, as computing servers are heterogeneous for both performance capacity and energy efficiency. It addresses that the heterogeneity of performance capacities can be worth to be considered when the performance is the main concern in systems.

References

1. Roy, A., Jingye, X., Chowdhury, H.: Multi-core processors: a new way forward and challenges in microelectronics. In: International Conference on ICM 2008, pp. 454–457 (2008)
2. Wang, L., Tao, J., von Laszewski, G.: Multicores in cloud computing: research challenges for applications. J. Comput. **5**(6), 958–964 (2010)
3. Nesmachnow, S., Dorronsoro, B., Pecero, J., Bouvry, P.: Energy-aware scheduling on multicore heterogeneous grid computing systems. J. Grid Comput. **11**, 1–28 (2013). Springer, Netherlands
4. NVIDIA: The Benefits of Multiple CPU Cores in Mobile Devices (2011)
5. Sanati, B., Cheng, A.M.K.: maximizing job benefits on multiprocessor systems using a greedy algorithm. SIGBED Rev. **5**, 3:1–3:4 (2008)
6. Jeet, R., Garg, U.: selective scheduling based on number of processor cores for parallel processing. Int. J. Sci. Res. (IJSR) **4**, 1666–1669 (2015)
7. Papazachos, Z.C., Karatza, H.D.: Gang scheduling in multi-core clusters implementing migrations. Future Gener. Comput. Syst. **27**, 1153–1165 (2011)
8. Xuan, T.T., Do, T.V.: Job scheduling in a computational cluster with multicore processors. In: Nguyen, T.B., Do, T.V., Le Thi, H.A., Nguyen, N.T. (eds.) Advanced Computational Methods for Knowledge Engineering. AISC, vol. 453, pp. 75–84. Springer, Heidelberg (2016). doi:10.1007/978-3-319-38884-7_6
9. Grochowski, E., Ronen, R., Shen, J., Wang, P.: Best of both latency and throughput. In: Proceedings of IEEE International Conference on Computer Design: VLSI in Computers and Processors ICCD 2004, pp. 236–243 (2004)
10. Ellision, B., Minas, L.: Energy Efficiency for Information Technology: How to Reduce Power Consumption in Servers and Data Centers. Intel press, Oakland (2009)
11. Qi, X., Zhu, D.-K.: Energy efficient block-partitioned multicore processors for parallel applications. J. Comput. Sci. Technol. **26**(3), 418–433 (2011)
12. NVIDIA: Variable SMP - A Multi-Core CPU Architecture for Low Power and High Performance, Technical report, NVIDIA's Project Kal-El, NVIDIA Corporation (2012)
13. Shieh, W.-Y., Pong, C.-C.: Energy and transition-aware runtime task scheduling for multicore processors. J. Parallel Distrib. Comput. **73**, 1225–1238 (2013)
14. Lammie, M., Brenner, P., Thain, D.: Scheduling grid workloads on multicor clusters to minimize energy and maximize performance. In: 2009 10th IEEE/ACM International Conference on Grid Computing, pp. 145–152, October 2009
15. Utrera, G., Tabik, S., Corbalan, J., Labarta, J.: A job scheduling approach for multi-core clusters based on virtual malleability. In: Kaklamanis, C., Papatheodorou, T., Spirakis, P.G. (eds.) Euro-Par 2012. LNCS, vol. 7484, pp. 191–203. Springer, Heidelberg (2012). doi:10.1007/978-3-642-32820-6_20
16. Zikos, S., Karatza, H.D.: Performance and energy aware cluster-level scheduling of compute-intensive jobs with unknown service times. Simul. Model. Pract. Theor. **19**(1), 239–250 (2011)
17. Do, T.V., Vu, B.T., Tran, X.T., Nguyen, A.P.: A generalized model for investigating scheduling schemes in computational clusters. Simul. Model. Pract. Theor. **37**, 30–42 (2013)
18. Standard Performance Evaluation Corporation. http://www.spec.org/
19. SPEC, Acer Incorporated Acer $AR380F2$, July 2012. http://www.spec.org/power_ssj2008/results/res2012q3/power_ssj2008-20120525-00481.html

20. SPEC, Fujitsu FUJITSU Server PRIMERGY tx1330 m1, July 2015. https://www.spec.org/power_ssj2008/results/res2015q1/power_ssj2008-20150116-00684.html
21. SPEC, Acer Incorporated Altos r360 f2, July 2013. https://www.spec.org/power_ssj2008/results/res2013q3/power_ssj2008-20130619-00617.html
22. SPEC, Fujitsu FUJITSU Server PRIMERGY tx1330 m2, January 2016. https://www.spec.org/power_ssj2008/results/res2016q1/power_ssj2008-20151214-00707.html

Secured-OFS: A Novel OpenFlow Switch Architecture with Integrated Security Functions

Bao Ho(✉), Quoc Nguyen, Cuong Pham-Quoc, and Tran Ngoc Thinh

Ho Chi Minh City University of Technology, Vietnam National University - HCMC, Ho Chi Minh City, Vietnam
{7140219,51102795,cuongpham,tnthinh}@hcmut.edu.vn

Abstract. Although OpenFlow network protocol is a promising network approach with many advantages compared to traditional network approaches, it still suffers from network attacks. In this paper, we propose a novel architecture for an OpenFlow-based switch with associated multiple network security techniques, so-called Secured-OFS. The proposed Secured-OFS can not only function as a switch following the OpenFlow protocol but also help protect a network against many attack types. We implement the first FPGA-based prototype version of our proposed Secured-OFS using a Xilinx Virtex 5 xc5vtx240t device. In this first prototype version, we integrate two different DDoS defense techniques, Hop-Count Filtering and Port Ingress/Egress Filtering. The experimental results show that the switch not only fulfills the OpenFlow protocol but also be able to defense against DDoS attacks. The system achieves a maximum throughput at 19.729 Gbps while a 100 % DDoS attack detection rate is obtained.

Keywords: Software defined networking · OpenFlow network · Network security

1 Introduction

In the last decades, SDN [1] has been considered as a promising paradigm to manage and configure computer networks through a high-level abstraction. Compared to the traditional approach where computer networks are configured manually, the SDN approach has many benefits such as centralization control and monitoring, simple hardware devices, and high virtualization. The SDN architecture decouples network control from forwarding functions so that network control becomes programmable. The network control includes *controllers* programmed by network administrators through software interfaces. Each controller is responsible for handling several *forwarding devices* behaving forwarding functions. Those forwarding devices route network packets from a source node to a particular destination node according to network configuration.

As the most well-known instance of SDN, OpenFlow [2] is not only a quite popular implementation in academia but also an industry standard of SDN [3].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_57

Based on the architecture of SDN, the OpenFlow network architecture also decouples network control function from the forwarding functions. Therefore, the OpenFlow network takes all the advantages of the SDN paradigm. Moreover, by optimizing elements such as controllers and forwarding devices, the OpenFlow network can be implemented as a software program or used as hardware platforms.

However, many security issues exist in both the SDN and OpenFlow network architectures. Research in [4] presents seven different threats in a SDN network which attackers can exploit to attack the network. OpenFlow networks also have some security threats that should be considered carefully. In the literature, there are some proposals to defend against possible attacks [5–7]. However, these approaches are only deployed as software programs. Moreover, research in the literature mainly focuses on optimizing controllers in OpenFlow networks [8,9]. There are still many open issues with forwarding devices. With the rapid progress in network services and network speed, high-performance and secure forwarding devices in OpenFlows networks is an essential demand.

With the fast increasing in the number of network attacks, hardware-based network defense plays an important role of a successful cyber-security strategy. In this paper, we propose a Secured-OpenFlow switch (Secured-OFS) with associated security functions. These Secured-OFSes can operate as forwarding devices in OpenFlow networks. We implement the first prototype version of Secured-OFS on the NetFGPA-10G [10] board which contains a Xilinx Virtex 5 xc5vtx240t FPGA device. The experimental results show that the system achieves high accuracy forwarding with nearly 100 % while the detection rate reaches to 100 % for two cases of IP Spoofing. Besides, the forwarding services time is 0.36 μs and 9.36 μs for the minimum and maximum packet respectively. That leads to the performance of total system reaching 9.859 Gbps in half-duplex and 19.718 Gbps in full-duplex mode. The main contributions of our paper can be summarized as follow.

- We propose a Secured-OFS with integrated security functions. To the best of our knowledge, this is the first OpenFlow switch that can not only route network packets according to the OpenFlow protocol but also defend against network attacks.
- We present our first prototype Secured-OFS using the NetFPGA-10G board which integrates two different DDoS defense mechanisms including Hop-Count Filtering and Port Ingress/Egress Filtering. The first prototype can work at up to 108.711 MHz and achieves a 99.1 % detection rate.

The rest of the paper is organized as follow. Section 2 presents the proposed architecture of Secured-OFS. Section 3 introduces our FPGA-based first prototype version of the proposed Secured-OFS. We analyze our experiments in Sect. 4. Finally, Sect. 5 concludes the paper and introduces future work.

2 Secured-OFS Architecture

In this section, we present our proposed Secured-OFS architecture. Our Secured-OFS can not only operate as an OpenFlow protocol-based switch but also a security device to defense against network attacks.

Figure 1 illustrates our proposed Secured-OFS architecture. The proposed Secured-OFS consists of three different components named *Ingress*, *Egress*, and *Engine*. The Ingress component is responsible for receiving incoming packets from input ports and forwarding to the Engine component for processing. Finally, those packets are routed to corresponding output ports of the Egress component.

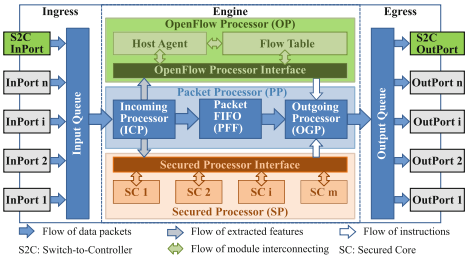


Fig. 1. Secured-OFS architecture

2.1 Ingress Component

The Ingress component includes one input queue, multiple data input ports (*InPort i*), and one control input port (*S2C-InPort*). All incoming packets from these input ports are collected and stored into buffers. Input Queue sequentially selects a packet from the buffer and forwards to the *Engine* component for processing. Input Queue can be configured on the fly so that packets from buffers are selected based on a specific strategy such as Round Robin or input port priorities. Configuration data is sent to Input Queue through the S2C-InPort.

S2C-InPort is the means of communication between a controller and the Secured-OFS. In other words, the corresponding controller sends configuration data through this port to handle the Secured-OFS according to the OpenFlow protocol. Compared to data input ports, this S2C-InPort has a higher priority, i.e., Input Queue selects packets coming from this port to send to the Engine component whenever there is any existence packet in the buffer of S2C-InPort regardless strategies used to select packets at Input Queue.

2.2 Egress Component

In contrast to the Ingress component, Egress consists of an output queue, several data output ports (*OutPort i*), and one control output port (*S2C-OutPort*). A packet after being processed by the Engine component is forwarded to the Output Queue. Regarding to routing information of the packet, Output Queue

sends it to a corresponding data output port. However, following the OpenFlow protocol, there are some cases in which a packet cannot be routed to any data output port due to the lack of information. In those cases, the packet is forwarded to the controller associating with the Secured-OFS through the S2C-OutPort so that the controller can update the Secured-OFS.

2.3 Engine Component

In our proposed architecture, the Engine component plays the most important role. It has much functionality than the works in [11–13]. The component processes an incoming packet that comes from the Ingress component following both OpenFlow protocol and the implementation of network security mechanisms. After processing a packet, the Engine component sends it to the Egress component so that the packet is forwarded to specific destination or the corresponding controller. This component consists of three different processors. Those are OpenFlow Processor (OP), Packet Processor (PP), and Secured Processor (SP).

OpenFlow Processor. The OP consists of a *Host Agent* module, a *Flow Table*, and an interface to communicate with PP. A packet can come to the Secured-OFS through two different port types, the data ports and the control port (S2C-InPort). A packet coming to the Secured-OFS through the control port is a packet generated by the associated controller. This packet contains an instruction that the controller uses to handle the Secured-OFS such as updating Flow Table or modifying a packet header. The Host Agent module is responsible for receiving control packets, executing instructions, and sending feedback to the controller if required when a control packet comes to the Secured-OFS. When a data packet is arriving at the Secured-OFS, it should be scanned by SP to defend against network attacks and processed by this OP following the OpenFlow protocol. The Host Agent module analyzes the data packet and retrieves the Flow Table to extract the corresponding actions associated with the packet. If an action for the packet is found, it will be sent to the PP. Otherwise, OP requests PP to send the packet to the associated controller. To support the OpenFlow protocol, the Flow Table module is needed to store OpenFlow-based actions [14] such as dropping, updating the header fields, or forwarding a packet to a destination output port. Figure 2 illustrates a segment of Flow Table.

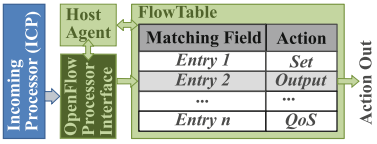


Fig. 2. The Flow Table architecture

Packet Processor. PP consists of three different modules the *Incoming Packet Processing*, the *Packet FIFO*, and the *Outgoing Packet Processing*. It is used to decode an incoming packet into different fields such as header field, and payload field at first. Depending on input port types, PP processes incoming packets in two different scenarios. In the first scenario, a coming packet through the control port is forwarded to the *Host Agent* module of OP without any processing by Incoming Packet Processing module.

In the second scenario, a packet coming through a data input port is analyzed by the Incoming Packet Processing module first. After that both OP and SP are activated simultaneously to process the packet. The header field of the packet is forwarded to the Host Agent module so that corresponding actions of this packet can be found. Depending on which network security mechanisms are used in SP, different fields of the packet are forwarded to SP for scanning. SP guarantees that the packet is a legitimate network packet. The whole packet is stored in the Packet FIFO to wait for decisions from both OP and SP. The Outgoing Packet Processing module starts processing the waiting packet in Packet FIFO immediately after receiving the final decisions from both SP and OP. If SP recognizes that the packet does not belong to a network attack, the taken actions from Flow Table are applied to the packet. Otherwise, the packet is destroyed immediately. In the case of OpenFlow-based actions for a particular packet cannot be found in Flow Table, i.e., the OP requests to send the packet to the associated controller, the Outgoing Packet Processing module forwards this packet to the S2C-OutPort.

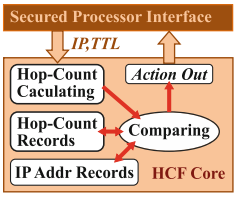
Secured Processor. The SP is responsible for guaranteeing that incoming packets do not belong to network attacks. SP includes a number of *Secured Cores* (SCs) and an interface for communication between the cores and PP. Each Secured Core implements a particular network defense mechanism to defend against a specific type of network attacks such as Anti-DDoS or Anti-Virus. Depending on characteristics of a network where the Secured-OFS is deployed, different security functions are chosen to implement in Secured Cores. Because each Secured Core performs one specific security function, a packet is considered as an illegal packet if there is at least one core alerts the packet is illegal. In this case, a drop signal is issued to PP. Otherwise, when all Secured Cores vote that the packet is legal, a pass signal is sent to the PP.

3 FPGA-Based Secured-OFS Prototype

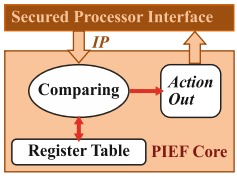
The previous section shows our proposed Secured-OFS architecture. The architecture can be developed by using many different technologies such as FPGA or ASIC. The proposed architecture also can be deployed using multicore systems where each processor in the Engine component can be implemented by a computing core. In this section, we present our first FPGA-based Secured-OFS prototype using the proposed architecture in the previous section.

In this first prototype version, we decide to build two well-known DDoS countering mechanisms, the Hop-Count Filtering (HCF) and the Port Ingress-Egress Filtering (PIEF), in SP because DDoS attacks have become one of the primary cyber-security threats [15,16]. These kinds of attack are attempts to make a computer resource (i.e. website, e-mail, VoIP, or a whole network) unavailable to its intended users. In the final quarter of 2015, the number of DDoS attacks has hit a new record [17].

The Hop-Count Filtering core comprises four main modules called *Hop-Count Calculating*, *Hop-Count Records*, *IP Add Records*, and *Comparing*. Figure 3a presents the architecture of the HCF secured core in this work. When the core receives a source IP address and its final Time-to-Live (TTL) value from PP through Secured Processor Interface, Hop-Count Calculating computes the Hop-Count value for the packet. Then the Comparing module look-ups IP address and stored Hop-Count value of the packet using both the Hop-Count Records and IP Addr Record module. If the calculated Hop-Count value is matched to the stored Hop-Count value, the packet is considered as the legitimate packet. In this case, a *bypass* signal is returned to PP. In the case of the calculated Hop-Count value is different from the stored Hop-Count value, the packet is classified to a DDoS attack. A *drop* signal is issued to alert PP. If this is the first time of a packet coming from this source IP address that has come to the switch, both IP address and TTL value of the packet are stored into Hop-Count Records and the IP Addr Record module. However, this is the main drawback of the Hop-Count defense mechanisms because IP address and TTL value of a packet can be modified while it is traveling in over networks. Therefore, in this work, we consider integrating multiple DDoS countermeasure techniques into a hardware system to improve both protection efficiency and system performance.



(a) HCF Core Architecture



(b) PIEF Core Architecture

Fig. 3. The Secured Cores in our first prototype

Figure 3b depicts the architecture for the PIEF core in this paper. As shown in the figure, the PIEF core contains a Register Table module and a Comparing module. Register Table stores special IP address blocks that are not allowed to appear in networks [18]. The Comparing module compares IP addresses of incoming packets with stored IP addresses in Register Table. When the source IP address of the packet is sent to PIEF, PIEF searches the address in Register Table. If a *miss* signal is returned (i.e., no record was found in Register Table),

the packet is legitimate. Otherwise, the packet is illegitimate (i.e., a *hit* signal is returned). Based on this hit or miss signal, the PIEF core send *bypass* or *drop* signal to PP.

In this first FPGA-based prototype Secured-OFS, the proposed architecture with two DDoS countermeasure mechanisms is built in a single FPGA chip. Therefore, all the inter-components, as well as intra-component interconnects are implemented as on-chip point-to-point and bus-based interconnects. The point-to-point interconnect is very suitable for FPGA implementation because FPGA includes many wires for interconnect. Moreover, the most advantage of the point-to-point interconnect is high-performance because there is no any competition for communication through the point-to-point interconnect. The bus-based interconnect is low latency and area-efficiency. However, bus-based interconnect suffers from low overall communication performance.

We use embedded on-chip memory (Block RAM - BRAM) to build *Flow Table*, *Register Table* in the PIEF secure core, and both *IP Addr Records* and *Hop-Count Records* in the HCF secure core. BRAM allows memory accessing within exactly one cycle so that it helps improve the system performance because of frequent Flow Table accessing. However, BRAM still comprises a disadvantage that is the limitation of usable resources and interconnects. To overcome this disadvantage, external memory chip can be addressed for the prototype of next version.

4 Experiments

In this section, we present our experiments with the first FPGA-based prototype Secured-OFS. We also analyze the hardware resources usage as well as system performance of the proposed Secured-OFS.

4.1 Experimental Setup

In order to validate and estimate the system performance of the first prototype version of our Secured-OFS, we employ two NetFPGA-10G [10] boards to build a testing system. Figure 4 illustrates our testing model. The first NetFPGA-10G board is configured as our FPGA-based Secured-OFS with four data ports using SFP + (Small Form-Factor Pluggable) interface. The PCI Express connector of the board is used as control ports (S2CInPort and S2COutPort) to communicate with OpenFlow Controller software running on the host. The second NetFPGA-10G board is configured as a *Test Agent* in which we port the framework of Open Source Network Test (OSNT) [19]. OSNT including a Generator and a Monitor can be used to generate packets under some parameters and monitor incoming network packets. The Generator is used to generate not only legitimate packets but also attacking packets at the line-rate ($\approx$ 10Gbps/port) to send testing data to the input ports of our Secured-OFS system. All legitimate packets are switched to one of three output ports of Secured-OFS from *SFP+Port0-Out* to *SFP+Port2-Out* depending on the handling of the controller. Instead of dropping

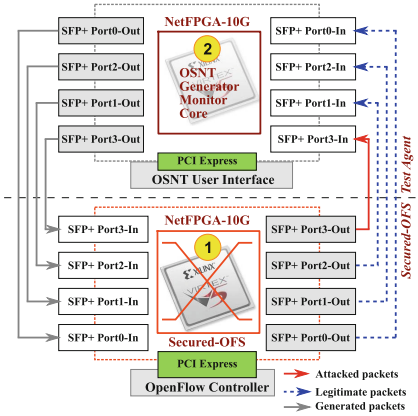


Fig. 4. The testing model of Secured-OFS

attacking packets as presented above, we configure the Secured-OFS to forward those packets to *SFP+Port3-Out* for doing statistic.

In our experiments, we conduct several test cases with different packets in term of size, header, and payload data. We measure and calculate the interval to process the *first-packet* coming from a specific source and the switching time of *non-first-packet* to validate the OpenFlow operation of the proposed Secured-OFS. When the first packet comes from a specific source, switching information of this packet does not exist in Flow Table. Following the OpenFlow protocol, the switch needs to forward it to the associated controller. We also collect data from the report of OSNT Monitor in which all packets are classified to check the capacity of DDoS protection of the secured cores such as detection rate, false positive rate, and false negative rate. Finally, we use several packet sizes to evaluate overall throughput of the proposed Secured-OFS.

4.2 Experiment Results

Table 1 summarizes hardware resources usage and synthesis frequency of the prototype Secured-OFS. Due to the fact that we use BRAM to implement Flow Table in OpenFlow Processor, IP Addr Records and Hop-Count Records in the HCF core, and Register Table in the PIEF core, the Secured-OFS consumes much more BRAM resource than other types. The first prototype version of our Secured-OFS can work at up to 108.711 MHz.

In performance evaluation experiment, generated packets belong to one of six types 64 bytes, 128 bytes, 256 bytes, 512 bytes, 1024 bytes, and 1500 bytes. The range from 64 bytes to 1500 bytes is valid for an Ethernet packet. We use these packet to test the OpenFlow protocol function of the switch in both half-duplex and full-duplex mode. The results of this performance evaluation experiment is shown in Fig. 5. In this chart, the horizontal axis indicates the *Packet size* in *bytes* while the vertical axis shows *Throughput* in *in Gbps* responding to each

Table 1. The hardware resource utilization of the system

Resources	Amount	Percentage
Look Up Tables (LUTs)	54770	36 %
Registers (FFs)	57424	38 %
Block RAM	204	62 %

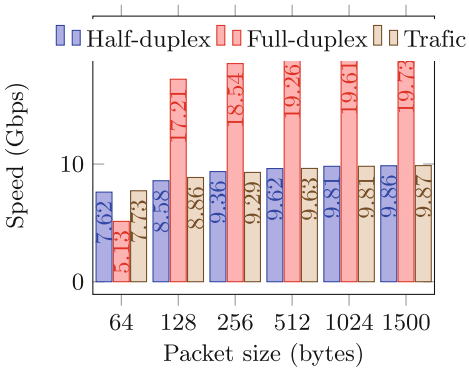


Fig. 5. Performance testing of Secured-OFS

packet types. In each packet type, the first column shows system throughput in the half-duplex mode. The second column depicts throughput in full-duplex mode. Finally, the last column presents OSNT packet generator speed. According to these experimental results, our first prototype system can achieve throughput by up to 9.859 Gbps in half-duplex mode and up to 19.729 Gbps in full-duplex mode.

In the switching time experiments, each flits in a packet (a packet is divided into a number of fixed-size flits) requires 15 cycles to be processed. Therefore, total forwarding time for a packet depends on the packet size. Table 2 presents processing time for each packet type. A 64-byte packet consumes 0.36 μs while a 1500-byte packet requires 9.1 μs to be processed. The proposed Secured-OFS is also validated in the case of first packet coming from a specific source arriving the switch. The experiments show that the functionality of our Secured-OFS is fully satisfied. Table 3 shows the detection rates, false positive rates, and false negative rates according to the packet sizes. The detection rates, false negative rates, and false positive rates are almost stable when the sizes of packets are changed. According to the table, the system can recognize all attacking packets although a 2.9 % false negative rate (legitimate packets are classified as attacking packets) occurs during the test.

Table 2. Timing testing of Secured-OFS

Packet size (bytes)	Forwarding time (μs)	First packet delay Time (μs)
64	0.36	0.08
128	0.7	0.14
256	1.38	0.26
512	2.76	0.51
1024	5.36	0.98
1500	9.36	1.78

Table 3. Detection rate of Secured-OFS *(in %)*

Type	<64	64–128	128–256	256–512	>512
Detection rate	97.1	97.6	98.4	98.6	99.1
False negative	0.0	0.0	0.0	0.0	0.0
False positive	2.9	2.4	1.6	1.4	0.9

5 Conclusion

In this paper, we propose a novel architecture for an OpenFlow switch which can defend against vulnerabilities from the network. This research contributes the novel framework which has high expansibility due to the Secured-OFS system based-on the combination with many secured functions. This architecture contains a processor able to merge many different secured core results into a final decision to against network intrusions. In the implementation, we integrate the HCF core and the PIEF core into Secured-OFS to achieve the high detection rate with DDoS attack. We have the plan to extend our Secured-OFS by integrating a dynamic reconfiguration module to this system. That helps the system become more flexible and efficient because the system could change secured cores to adapt to network attack types.

Acknowledgement. This research is funded by Ho Chi Minh City University of Technology (HCMUT) under grant number TSDH-2015-KHMT-52.

References

1. Goransson, P., Black, C.: Software Defined Networks: A Comprehensive Approach. Morgan Kaufmann, San Francisco (2014)
2. McKeown, N., Anderson, T., Balakrishnan, H., Parulkar, P., Peterson, L., Rexford, J., Shenker, S., Turner, J.: OpenFlow: enabling innovation in campus networks, vol. 38, pp. 69–74. ACM, New York (2008)
3. Gelberger, A., Yemini, N., Giladi, R.: Performance analysis of software-defined networking (SDN). In: 2013 IEEE 21st International Symposium on Modelling, Analysis and Simulation of Computer and Telecommunication Systems. IEEE Press, San Francisco (2013)

4. Kreutz, D., Ramos, F., Verissimo, P.: Towards secure and dependable software-defined networks. In: Proceedings of the Second ACM SIGCOMM Workshop on Hot Topics in Software Defined Networking, pp. 55–60. ACM, New York (2013)
5. Tootoonchian, A., Ganjali, Y.: HyperFlow: a distributed control plane for OpenFlow. In: Proceedings of the 2010 Internet Network Management Conference on Research on Enterprise Networking. USENIX Association, Berkeley (2010)
6. Braga, R., Mota, E., Passito, A.: Lightweight DDoS flooding attack detection using NOX/OpenFlow. In: 2010 IEEE 35th Conference on Local Computer Networks (LCN), pp. 408–415. IEEE, Denver (2010)
7. Shin, S., Song, Y., Lee, T., Lee, S., Chung, J., Porras, P., Yegneswaran, V., Noh, J., Kang, B.B.: Rosemary: a robust, secure, and high-performance network operating system. In: Proceedings of the 2014 ACM SIGSAC Conference on Computer and Communications Security, pp. 78–89. ACM, New York (2014)
8. Hu, Y., Wang, W., Gong, X., Que, X., Cheng, S.: Balanceflow: controller load balancing for openflow networks. In: 2012 IEEE 2nd International Conference on Cloud Computing and Intelligent Systems (CCIS), vol. 2, pp. 780–785. IEEE, New York (2012)
9. Chen, X., Zheo, B., Ma, S., Chen, C., Hu, D., Zhou, W., Zhu, Z.: Leveraging master-slave OpenFlow controller arrangement to improve control plane resiliency in SD-EONs. In: Optics Express, vol. 23, no. 6, pp. 7550–7558. Optical Society of America (2015)
10. NetFPGA 10G. http://netfpga.org/site/#/systems/3netfpga-10g/details/
11. Naous, J., Erickson, D., Covington, G.A., Yang, S., Appenzeller, G., McKeown, N.: Implementing an OpenFlow switch on the NetFPGA platform. In: Proceedings of the 4th ACM/IEEE Symposium on Architectures for Networking and Communications Systems, pp. 1–9. ACM, New York (2008)
12. OpenFlow Implementation on NetFPGA-10G. https://docs.google.com/document/d/1ZwHXQZocKwQls6Ted8VZO8h9MjBtu9WxV2fAY44eOgE/edit
13. Naous, J., Erickson, D., Covington, G.A., Yang, S., Appenzeller, G., McKeown, N.: From 1G to 10G: code reuse in action. In: Proceedings of High performance and programmable networking, pp. 31–38. ACM, New York (2013)
14. Suh, M., Park, S.H., Lee, B., Yang, S.: Building firewall over the software-defined network controller. In: 16th International Conference on Advanced Communication Technology, pp. 744–748. IEEE, Pyeongchang (2014)
15. Zargar, S.T., Joshi, J., Tipper, D.: A survey of defense mechanisms against distributed denial of service (DDoS) flooding attacks. In: IEEE Communications Surveys & Tutorials, vol. 15, no. 4, pp. 2046–2069. IEEE (2013)
16. Yan, Q., Yu, F.R., Gong, Q., Li, J.: Software-defined networking (SDN) and distributed denial of service (DDoS) attacks in cloud computing environments: a survey, some research issues, and challenges. In: IEEE Communications Surveys & Tutorials, vol. 18, pp. 602–622. IEEE (2016)
17. State of the Internet Report (2016). https://www.stateoftheinternet.com/resources-cloud-security-2015-Q4-web-security-report.html
18. Cotton, M., Vegoda, L.: Special Use IPv4 Addresses. Technical report, RFC-57 (2010)
19. The open source network tester. http://osnt.org/

Semi-supervised Clustering in Fuzzy Min-Max Neural Network

Dinh Minh Vu[1(✉)], Viet Hai Nguyen[1(✉)], and Ba Dung Le[2(✉)]

[1] Thai Nguyen Industry College, Thai Nguyen, Viet Nam
minhvd.itvn@gmail.com, nvhai81@gmail.com
[2] Hung Yen University of Technology and Education, Hung Yen, Viet Nam
l_bdung@yahoo.com

Abstract. The Fuzzy Min max Neural Network (FMNN) developed by Simpson is defined as a neural network that forms hyperboxes for classification and prediction. This paper proposes an improvement in learning algorithm in FMNN using semi-supervised clustering method, called SS-FMM. The proposed model combines the advantages of supervised learning and those of unsupervised learning. Labeled a part of data is the additional information that is used in this semi-supervised clustering method. For evaluation purpose, this algorithm is implemented on two datasets including Shape sets from CS and Thyorid disease from UCI. A part from that, in this paper, some related algorithms in FMNN are also setup on these datasets in order to compare the accuracy with proposed algorithm. The test results show that the novel algorithm has the better performance.

Keywords: Semi-supervised · Clustering · Fuzzy min-max · Neural networks

1 Introduction

From the fuzzy set theory [15] launched by Zadeh, there are many researches on the identification and classification with fuzzy sets. In particular, many scientists paid attention to the fusion of fuzzy logic and neural networks to develop the intelligent systems [4,12,14]. This leads to the aggregation of fuzzy reasoning in handling uncertain information and the learning ability of fuzzy neural networks [3,13,16].

Based on the advantages of the combination between fuzzy logic and neural networks, Simpson has proposed a new approach called "Fuzzy Min-Max Neural Network" (FMNN) using an aggregation of hyperbox sets [10,11]. This model combines neural network and fuzzy min-max (FMM) theory in order to solve the classification and clustering problems. The FMNN model using the data-based method is the reinforcement learning neural network that is able to carry out the problem with big data sets [6]. Reinforcement learning is an effective technique in knowledge discovery because it allows to reuse and to add more information

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_58

in a single pass through [5]. Another advantage of reinforcement learning is all training data can be used immediately for learning process and in afterward iterations without retraining.

In FMNN, each of the hyperboxes is considered as a subspace in the n-dimensional space. FMNN based on the synthetic from this hyperbox [1] is used to determine and to limit corresponding subspace. The hyperbox is completely defined by its minimum and maximum points and the membership functions corresponding to data points in this hyperbox. Each pattern is classified based on the value of its membership function. Simpson used two main training strategies for pattern classification procedures in FMNN including supervised learning [10] and unsupervised learning [11].

Based on the original FMNN model, Gabrys *et al.* [2] proposed a general fuzzy min-max neural network (GFMM). This is a generalization and extension of the fuzzy min-max clustering and classification algorithms. In this paper, GFMM model allowed processing both fuzzy (hyperboxes in pattern space) and crisp (points in pattern space) input patterns. Moreover, the training of the GFMM neural network is very fast and, as long as there are no identical data belonging to two different classes, the recognition rate for training data is 100 %. This proposal still has some limitations. Firstly, its performance strongly depends on the characteristics of the training and test data. Secondly, it meets the challenge in dealing with elongated and rotated clusters of hyperelipsoidal data.

Quteishat and Lim [9] introduced a modification of FMM network in an attempt to improve its classification performance in situations when large hyperboxes are formed by the network. However, this modification is used to solve the problem when the number of hyperboxes created by the network is small.

A new fuzzy Min-Max classifier was proposed using modified compensatory neurons (FMCN) by Nandedkar and Biswas [8]. The main advantage of this model is that its performance is less dependent on the initialization of expansion coefficient. In this study, FMCN is also compared with original fuzzy min-max neural network classifier and GFMM.

Mohammed and Lim [7] proposed an enhanced fuzzy min-max (EFMM) network for pattern classification. The new network includes three heuristic rules to enhance the learning algorithm of FMM. In this paper, the authors also showed that EFMM network archives better results than FMM-based models, support vector machine-based, Bayesian-based, decision tree-based, fuzzy-based, and neural-based classifiers.

In this paper, we propose a novel learning algorithm in FMNN using semi-supervised learning method (SS-FMM). This improved algorithm combines the advantages of supervised and unsupervised learning in FMM. During training process, the labeled patterns are taken first following by unlabeled patterns. The learning algorithm labels unlabeled patterns in the process of learning and clustering. This algorithm is also implemented on two datasets consisting of Shape sets from CS and Thyroid dataset from UCI. A part from evaluating the proposed algorithm, this paper also gives the comparison to other available methods such as general Fuzzy Min-Max Neural Network (GFMM) and Enhance

Fuzzy Min-Max Neural Network (EFMM). Experimental results show that the accuracy of the our algorithm is better than other related methods. Moreover, our algorithm uses a small amount of the labeled patterns.

The main contributions of this paper are: (i) Proposing a novel algorithm using semi-supervised clustering method in fuzzy min-max neural network; (ii) Implementing this algorithm on Shape sets and Thyroid disease set in order to evaluate the performance of new algorithm; (iii) Comparing the accuracy and number of hyperboxes between proposed algorithm and other related methods. The experimental results show that our new algorithm has the higher accuracy with less hyperboxes than others.

The rest of this paper is organized as follows: Sect. 2 gives the overview of the development of the FMNN. Section 3 presents the semi-supervised learning in FMNN. Section 4 shows the experimental results in order to evaluate the performance of proposed method and to compare new method with other methods. Conclusion and further researches are presented in Sect. 5.

2 Fuzzy Min Max Neural Network

2.1 Hyperbox Membership Function

A hyperbox [10] is a simple geometrical structure (Fig. 1). An n-dimensional hyperbox can be defined by its min point and max point. The size of each dimension of a hyperbox is a value ranging from 0 to 1. The pattern space denoted by I^n is a hyperbox with size of the dimension as 1. The membership function of a data point shows the degree of that point belonging to the hyperbox.

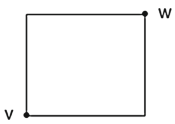


Fig. 1. A min-max hyperbox in 2D.

Denote N is the number of hyperboxes in the FMNN. A j^{th} hyperbox is defined as follow:

$$B_j = \{A_h, V_j, W_j, b_i(A_h, V_j, W_j)\}, j = 1, ..., N \tag{1}$$

for all $h = 1, 2, \ldots, m$ where $A_h = (a_{h1}, a_{h2}, \ldots, a_{hn}) \in I^n$ is the h^{th} pattern in the data set, $V_j = (v_{j1}, v_{j2}, \ldots, v_{jn})$ is min point for the j^{th} hyperbox B_j, $W_j = (w_{j1}, w_{j2}, \ldots, w_{jn})$ is max point for the j^{th} hyperbox B_j, and the membership function for the j^{th} hyperbox is $b_j\,(A_h, V_j, W_j)$ with the $0 \leq b_j\,(A_h, V_j, W_j) \leq 1$.

The membership function measures the degree that the hth input pattern Ah falls within the hyperbox. It is defined by Eq. (2) below.

$$b_j = \frac{1}{n}\sum_{i=1}^{n}\left[1 - f\left(a_{hi} - w_{ji}, \gamma\right) - f\left(v_{ji} - a_{hi}, \gamma\right)\right] \tag{2}$$

where $f\left(x, y\right)$ is the two-parameter ramp threshold function:

$$f(x,y) = \begin{cases} 1 & if\, xy > 1 \\ xy & if\, 0 \le xy \le 1 \\ 0 & if\, xy < 0 \end{cases}$$

The parameter γ the sensitivity parameter that makes the membership values decrease rapidly when an input pattern is separated from the hyperbox.

2.2 Fuzzy Min Max Learning Algorithm

Assumption that the training set D consists of the m patterns and denote $A_h \in D$ is a the h^{th} pattern $(h = 1, 2, \ldots, m)$. Based on the input patterns, FMNN creates a number of hyperboxes. The learning process starts by selecting an pattern A_h and finding the closest hyperbox to selected pattern that hyperbox can expand to cover the pattern. If such hyperbox can not be found, a new hyperbox is formed and added to the system. This process allows new clusters to be added without retraining.

One of the problems of hyperbox expansion is the overlapping which causes ambiguity. For each pattern belonging to the overlapping area, the value of its membership function always equals to 1 (for every overlapping hyperbox). The FMNN utilizes a contraction process to eliminate any hyperbox overlap.

The detailed learning steps are as follows.

Step 1 (expansion): Given a pattern A_h, the hyperbox B_j that is selected for the expansion process must satisfy two criteria. One is the value of membership function of given pattern to the hyperbox is the highest. Another is the maximum size of the hyperbox is bounded by θ where $\theta \in [0, 1]$ is a threshold defined by user. This means the following constraint must be met:

$$\sum_{i=1}^{n}\left(max\left(w_{ji} - A_{hi}\right) - min\left(v_{ji} - A_{hi}\right)\right) \le n\theta \tag{3}$$

If (3) is satisfied, the min and max points of the hyperbox are adjusted by using the Eqs. (4) and (5):

$$v_{ji}^{new} = min\left(v_{ji}^{old}, a_{hi}\right), \forall i = 1, 2, \ldots n \tag{4}$$

$$w_{ji}^{new} = max\left(w_{ji}^{old}, a_{hi}\right), \forall i = 1, 2, \ldots n \tag{5}$$

If there is no expansion from all hyperboxes, then create a new hyperbox and the min-max points of this hyperbox are set by (6).

$$V = W = A_h \tag{6}$$

Step 2 (overlapping test): Assumption that the hyperbox B_k is expanded in the previous step. To determine overlap, a dimension-by-dimension comparison is conducted between B_k and the others. Let B_j and $B_j \neq B_k$, the expansion creates an overlap between B_j and B_k if one of the four cases is satisfied on each dimension of the hyperbox:

– Case 1: max of B_j overlaps min of B_k:

$$v_{ji} < v_{ki} < w_{ji} < w_{ki} \tag{7}$$

– Case 2: min of B_j overlaps max of B_k:

$$v_{ki} < v_{ji} < w_{ki} < w_{ji} \tag{8}$$

– Case 3: B_k contained within B_j:

$$v_{ji} < v_{ki} \leq w_{ki} < w_{ji} \tag{9}$$

– Case 4: B_j contained within B_k:

$$v_{ki} < v_{ji} \leq w_{ji} < w_{ki} \tag{10}$$

Step 3 (contraction): if the hyperboxes B_j and B_k are overlapped one dimension of hyperboxes, the overlapping is eliminated on that dimension based on four discussed cases. The overlapping hyperboxes are contracted as follows:

– Case 1: if $v_{ji} < v_{ki} < w_{ji} < w_{ki}$ then:

$$v_{ki}^{new} = w_{ji}^{new} = (v_{ki}^{old} + w_{ji}^{old})/2 \tag{11}$$

– Case 2: if $v_{ki} < v_{ji} < w_{ki} < w_{ji}$ then:

$$v_{ji}^{new} = w_{ki}^{new} = (v_{ji}^{old} + w_{ki}^{old})/2 \tag{12}$$

– Case 3: if $v_{ji} < v_{ki} \leq w_{ki} < w_{ji}$ then:

if $w_{ki} - v_{ji} < w_{ji} - v_{ki}$ then:

$$v_{ji}^{new} = w_{ki}^{old} \tag{13}$$

otherwise:

$$w_{ji}^{new} = v_{ki}^{old} \tag{14}$$

– Case 4: if $v_{ki} < v_{ji} \leq w_{ji} < w_{ki}$ then:

if $w_{ki} - v_{ji} < w_{ji} - v_{ki}$ then:

$$v_{ki}^{new} = w_{ji}^{old} \tag{15}$$

otherwise:

$$w_{ki}^{new} = v_{ji}^{old} \tag{16}$$

Steps 1–3 are repeated for each pattern in the data set until cluster stability is achieved.

2.3 Fuzzy Min Max Neural Network

The FMNN clustering can be implemented as a two-layer neural network (Fig. 2.a). The input layer, F_A consists of n processing elements (PEs) (one PE corresponding to one dimension of the pattern) and the output layer, F_B consists of m PEs (corresponding to m clusters). Each the neuron in input layer is connected to jth neuron in output layer via a dual weight consisting of two vectors V_j and W_j (Fig. 2.b).

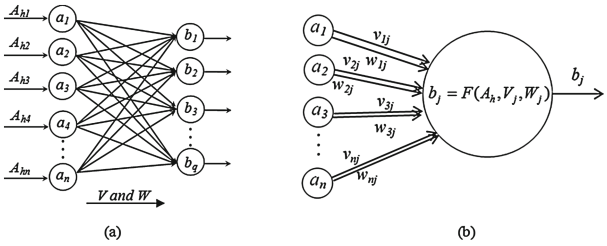


Fig. 2. (a) FMNN Architecture and (b) The structure of a neuron b_j.

3 FMNN Using Semi-supervised Learning

3.1 General Diagram of SS-FMM Algorithm

This section describes the semi-supervised learning algorithm used in clustering problem based on labeling a part of data.

The input data set D of this algorithm includes the pairs of $\{A_h, d_h\}$ where $d_h \in \{0, 1, \ldots, p\}$ is the cluster index ($d_h = 0$ for unlabeled patterns), $A_h = (a_{h1}, a_{h2}, \ldots, a_{hn})$ is a the h^{th} pattern ($h = 1, 2, \ldots, m$) of the D set. The learning algorithm to assign the label for the unlabeled patterns and clusters for the training samples. The general diagram of this network is presented as in Fig. 3 below.

From Fig. 3, input is a pattern from the dataset. In which, the labeled input patterns are taken first, the unlabeled input patterns were taken after. With labeled patterns, the SS-FMM algorithm will generate different hyperboxes. This process is similar to the learning process of FMNN proposed by Simpson. On other hand, with unlabeled patterns, SS-FMM algorithm finds the closest hyperbox by using membership function in (2) and the this hyperbox ought to satisfy the constraint in (3). When the needed hyperbox is selected, its min-max points are adjusted as in (4), (5) respectively. Consequently, input patterns are labeled. In the case of none of hyperboxes satisfying constraint (3), a new hyperbox is created by initializing the min-max points as in Eq. (6). In this case, the input pattern is labeled by label of the nearest hyperbox. Otherwise, a new hyperbox is created if the value of membership function that this pattern belongs to the

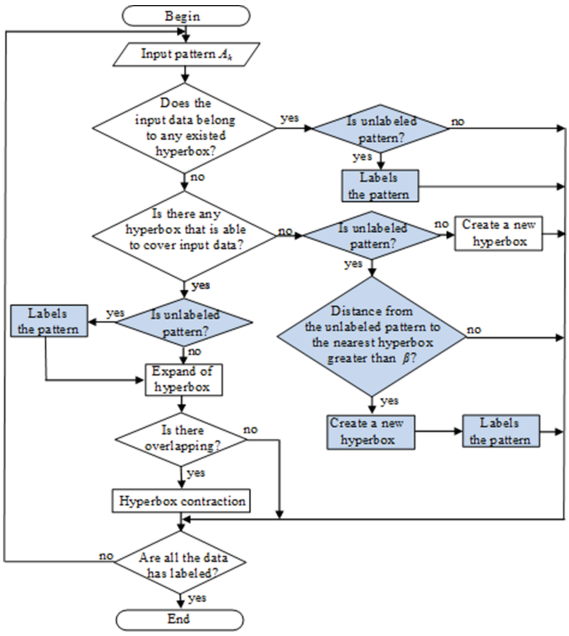


Fig. 3. General diagram of SS-FMM algorithm.

nearest hyperbox is greater than a threshold β where β is provided by user and it is considered as minimum limitation. In contrast, if that value is smaller than β, none of new hyperbox is generated. In this case, the input pattern is skipped and it will be considered as a new one in the next data browser. Hyperbox overlapping test between hyperboxes is performed by using Eqs. (7)–(10). If there is an overlap between the hyperboxes, the size of overlapping hyperboxes B_j is adjusted according to the V_j and W_j by applying one of four mentioned cases.

The SS-FMM algorithm stops when all the unlabeled patterns are labeled.

3.2 Main Steps of SS-FMM

Because this algorithm performs similar to the original FMNN with labeled patterns. From the diagram above, we just concentrate to process unlabeled patterns. The proposed algorithm is described in detail as below.

1. Give an unlabeled pattern
2. Find a hyperbox B_j such that offering largest membership and satisfying the condition (3) following these instances:
   - If input data belong to any existed hyperbox then labels the pattern.
   - Else If there any hyperbox that is able to cover input data then adjust min-max points of hyperbox B_j as in (4), (5) then labels the pattern.
3. If no suitable B_j is found then

- If $b_j > \beta$ then create a new hyperbox with $V = W = A_h$ and labels the pattern.

4. Hyperbox overlapping test between hyperboxes is performed by using Eqs. (7)–(10).

If there is an overlap between the hyperboxes, the size of overlapping hyperboxes B_j is adjusted according to the V_j and W_j by applying one of four mentioned cases as (11)–(16).

4 Experiments and Results

The SS-FMM algorithm is implemented on two datasets Shape sets from CS and Thyroid dataset from UCI. Shape sets have four separated datasets in two-dimension space including Flame, Jain, Spiral, Aggregation. In which, Flame dataset consists of 240 patterns, Jain dataset consists of 373 patterns, Spiral dataset includes 312 patterns and Aggregation dataset has of 788 patterns. Each of these datasets has different shape. The experiment uses 20 % of the labeled patterns and 80 % remaining unlabeled patterns.

The Thyroid dataset contains information about patients with thyroid disease. This dataset includes 7200 patterns with 21 attributes divided into two groups. Group 1 consists of 6666 patterns without the disease and group 2 consists of 534 patterns with the thyroid disease. In our experiments, 70 % of the patterns is used for the training, and the rest of patterns is used for testing. The parameters are set as $\gamma = 10$ and $\beta = 0.9$.

While the original FMNN with unsupervised learning algorithm showed the results that are not good enough on each of four discussed Shape sets, SS-FMM with semi-supervised learning achieves much better results (results in Table 1).

Table 1. Statistics of results on the Shape sets.

Data	Hyperbox size	No. of hyperboxes	Accuracy (%)
Flame	0.03	19	99.3
Jain	0.06	15	100
Spiral	0.02	56	100
Aggregation	0.03	57	99.5

Visually, Fig. 4 shows the scatter plot data and hyperboxes created by using new algorithm. On Thyroid disease dataset, corresponding to 30 %, 40 %, 50 % and 60 % of the dataset were selected randomly for training and the rest of dataset for testing, the performance of GFMM, EFMM and SS-FMM is shown in Table 2 (the size of hyperbox is $\theta = 0.1$).

From this table, the accuracy of SS-FMM is higher than other methods. Moreover, the number of hyperboxes (NoH) created by using SS-FMM is smaller

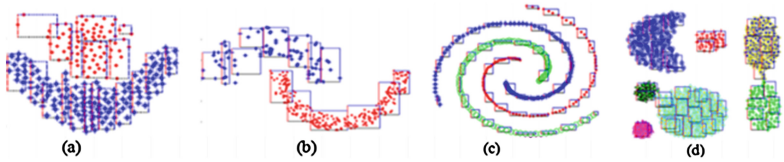


Fig. 4. Implemental results of SS-FMM on various datasets: (a) Flame; (b) Jain; (c) Spiral; (d) Aggregation

Table 2. Accuracy comparison on Thyroid data set among related methods.

Data	GFMM		EFMM		SS-FMM	
	Accuracy	NoH	Accuracy	NoH	Accuracy	NoH
30 %	91	127	93.7	147	94.3	127
40 %	92.3		94.1		94.5	
50 %	93.7		94.4		94.7	
60 %	94.4		94.7		95	

than this number when using EFMM. Comparing between SS-FMM and GFMM, the number of hyperboxes is equal but the accuracy of SS-FMM is a completely higher. These results show that, the performance of proposed algorithm is better than GFMM and EFMM.

5 Conclusion

In this paper, we concentrate to improve the learning method in FMNN. Firstly, based on original FMNN, a novel algorithm using semi-supervised clustering method in FMNN namely as SS-FMM is proposed. This algorithm combines the advantages of supervised and unsupervised learning in FMNN. Secondly, the proposed algorithm and GFMM and EFMM methods are illustrated on typical datasets. Thirdly, using the numerical results, the paper shows a better performance of new algorithm comparing with GFMM and EFMM methods in accuracy and number of hyperboxes. Besides, this algorithm only uses a small percentage of the labeled patterns from the datasets in whole training and testing progress. These are the contributions of this paper. Further works of this research can be investigated in the following ways: (i) Improve the computing rate in order to decrease the time requirement; (ii) Apply this algorithm on other datasets; (iii) Compare this algorithm with more other methods by various validity indices.

References

1. Alpern, B., Carter, L.: The hyperbox. In: Proceedings of the IEEE Conference on Visualization, Visualization 1991, pp. 33–139. IEEE Press, October 1991

2. Gabrys, B., Bargiela, A.: General fuzzy min-max neural network for clustering and classification. IEEE Trans. Neural Netw. **11**(3), 769–783 (2000)
3. Kosko, B.: Neural Networks and Fuzzy Systems: A Dynamical Systems Approach to Machine Intelligence/Book and Disk. Prentice Hall, Upper Saddle River (1992)
4. Lin, F.J., Shen, P.H.: Robust fuzzy neural network sliding-mode control for two-axis motion control system. IEEE Trans. Ind. Electron. **53**(4), 1209–1225 (2006). IEEE Press
5. Luo, C., Li, T., Chen, H., Liu, D.: Incremental approaches for updating approximations in set-valued ordered information systems. Knowl. Based Syst. **50**, 218–233 (2013)
6. Martínez-Rego, D., Fontenla-Romero, O., Alonso-Betanzos, A.: Nonlinear single layer neural network training algorithm for incremental, nonstationary and distributed learning scenarios. Pattern Recogn. **45**(12), 4536–4546 (2012)
7. Mohammed, M.F., Lim, C.P.: An enhanced fuzzy min-max neural network for pattern classification. IEEE Trans. Neural Netw. Learn. Syst. **26**(3), 417–429 (2015). IEEE Press
8. Nandedkar, A.V., Biswas, P.K.: A fuzzy min-max neural network classifier with compensatory neuron architecture. IEEE Trans. Neural Netw. **18**(1), 42–54 (2007)
9. Quteishat, A.M., Lim, C.P.: A modified fuzzy min-max neural network and its application to fault classification. In: Saad, A., Dahal, K., Sarfraz, M., Roy, R. (eds.) Soft Computing in Industrial Applications. ASC, vol. 39, pp. 179–188. Springer, Heidelberg (2007)
10. Simpson, P.K.: Fuzzy min-max neural networks. I. Classification. IEEE Trans. Neural Netw. **3**(5), 776–786 (1992)
11. Simpson, P.K.: Fuzzy min-max neural networks - Part 2: clustering. IEEE Trans. Fuzzy Syst. **1**(1), 32–45 (1993)
12. Wai, R.J., Lee, J.D.: Adaptive fuzzy-neural-network control for maglev transportation system. IEEE Trans. Neural Netw. **19**(1), 54–70 (2008)
13. Wang, Z., Zhang, H., Yu, W.: Robust stability of Cohen-Grossberg neural networks via state transmission matrix. IEEE Trans. Neural Netw. **20**(1), 169–174 (2009)
14. Yilmaz, S., Oysal, Y.: Fuzzy wavelet neural network models for prediction and identification of dynamical systems. IEEE Trans. Neural Netw. **21**(10), 1599–1609 (2010)
15. Zadeh, L.A.: Fuzzy sets. Inf. Control **8**(3), 338–353 (1965)
16. Zhang, H., Luo, Y., Liu, D.: Neural-network-based near-optimal control for a class of discrete-time affine nonlinear systems with control constraints. IEEE Trans. Neural Netw. **20**(9), 1490–1503 (2009)

Smart Lecture Room for Smart Campus Building Automation System

Kien Tran Pham Thai(✉)

Thai Nguyen University of Information and Communication Technology, Quyet Thang, Thai Nguyen, Vietnam
tptkien@ictu.edu.vn

Abstract. In the recent years, the terms of "Internet of Things" and "Smart Building Automation System" come into practical implementations from a hot topic for researches and developments. Beyond the concept of facilities or appliances in an apartment or office can be controlled remotely, as an automated entity, such the system is now required to be intelligent. The intelligence include an adaptive operation with context awareness, energy efficiency, and comfort living experience. However, it is difficult to realize such the complicated systems due to a gap between a high levels of abstraction in expectations to real devices. The purpose of this work is to realize a Smart Lecture Room, from a classical campus lecture room, as a complete solution for both building managers and users. During the development, the design process for this particular system is constructed. The realized system is then be analyzed to have better improvements in the future.

Keywords: Smart Lecture Room · Smart building management system · Smart cities · Real-time embedded system · MVC JavaFX Embedded system · Multitasking system · ZigBee two ways communication

1 Introduction

Current technological world is witnessing a new evolution of Internet of things from a hot trend to a common interest of both university research projects and market providers for practical applications. Among those applications, Building Automation System (BAS) standouts for not only capability of serving many applications but also providing possibilities to save significant amount of wasted energy.

In [1], the group of authors had mentioned HBAS term (Home and Building Automation System) where both systems is grouped together in one treatment.

In order to cope with the complexity of a building system, a new concept has been carried out by researchers by definitions of control zones [2]. In more details, rooms are categorized in zones for conditional optimization/control strategy for that unit area. The zone division approach is also discussed in [3] where the term of smart spaces is introduced.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_59

From [4,5], simple setups for building management system have presented with real time approach. The real time terms are again just the description of data collection methods instead of the application of actual RTOS.

In [6,7], Service Oriented Architecture (SOA) has been defined, migrated from industrial context to be in smart home or buildings. This promised to be a modern intelligent approach for system design due to its advantages. A SOA system must be able to perform both features: independently work as an autonomous system and cooperatively work with other systems. Consequently, the usage of native SOA will be a complicated and costly approach. As the results, there is a need for a better solution.

In [8], we have a practical description of the SOA in the Building Automation System. As revealed, the SOA devices are built with a configuration of both sensors and actuators and use wring bus as management protocol. In the paper, each controlled appliance has its services and all types of services are visualized.

Previously, ZigBee was mentioned as a candidate for wireless home communication network in [4,5].

In [11,12], there are several device settings in both hardware and software have been proposed. The device normally hosted by a Micro Controller (Arduino) with peripherals for sensing, displaying, actuating and communicating. In term of software, several Real Time Operating System (ScmRTOS, TinyOS) was installed. This construction allow the devices to meet the SOA concept and be a good reference for further design and development. Further investigations on communication technology are in [9,10,13], which have presented advantages of ZigBee technology.

This paper proposes a complete solution for realizing an actual Smart Lecture Room in real devices from defined abstract requirements. The Smart Lecture Room offers both automation features and smart saving features. The paper aim to provide material for generalization of the design process and metrics to measure future enhancements.

The rest of the article is organized as follows: Sect. 2 will describe the phases of design process from methodology to Pre-prototyping phase by sub-dividing the whole system into sub-systems and the realization phase of a complete system. The results of the work is discussed in Sect. 3. Finally, the analyses of the prototyped system are presented in Sect. 4.

2 Prototyping

2.1 Methodology

The adapted design flow starts with a definition of Design concept in the chosen scope of the Smart Lecture Room. The room is a typical room setting in university campus. Then, the Research phase will investigate the current situation of equipped facilities of the room for later formation of system requirements. At the next stage, the Subsystems distribution and design step is based on the divide and conquer strategy. This is where the whole complicated system is divided into

subsystems. After that, the phase of Subsystems prototyping and debugging for the subsystems takes place. At the end, the subsystems are integrated before tests and evaluations in Fig. 1.

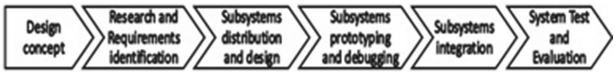


Fig. 1. Smart Lecture Room design process.

2.2 Pre-prototyping Phase

Design concept
The transformation of a room begins with an existing infrastructure of a typical lecture room Fig. 2. The infrastructure is then investigated, analyzed and added with automation features. The automation features allow users, from one centralized interface to interact or control the room itself. To equip the automated room with intelligence, more intelligent features are added. The intelligence enables the room to have its own decisions on its controlled objects. This is important to offer self-energy-saving when no presence of human control in the room. The objective of the system is not only to make a room to be more intelligent for users but also to connect all the smart rooms in a building in one management interface for building manager. These are the responsibilities of the management features.

Field investigation and problem identification
The investigation process takes place at Vrije Universiteit Brussel (VUB). Among the buildings, the D building is selected for providing a picture of actual settings in term of Dimension, Light system, Window ventilating and Safety and Security dimension (Fig. 3). Then, problems are stated.

Among the invested rooms, the conventional system reveals the disadvantages. The typical rooms manually use electrical components with no presence of automation at any centralized interface. Obviously, those regular electrical components have no inter-operation or communication at any level. The electricity energy can be wasted from the usage of the lighting system where users

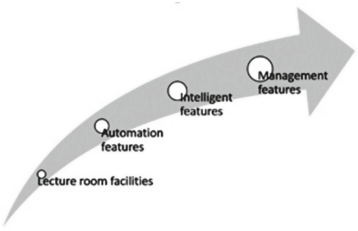


Fig. 2. The transformation of a conventional lecture room to a smart room.

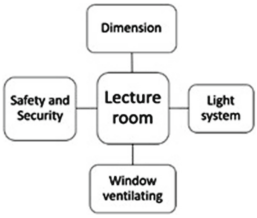


Fig. 3. The invested building D of VUB Campus with 4 factors for consideration in each of the rooms.

leave the room with the lights on until the building operator visits and shuts it down. In addition to that, door locks are maintained with keys information in another building, manually. The given rights for each key are limited in a hard written list of rooms to access and no presence of dynamical accesses. For example, to offer accesses to a certain room in certain times of the days or the year. Furthermore, the users have no information about their surrounded environment, which can be a informative knowledge or a major source of reference for control strategies. For the case when states of the doors or the windows need to be checked, for example, the users leave windows open after opening for fresh air, there is no solution to have the situation report automatically unless paying a visit. In hazard situation, when the windows are broken by either an intruder or hard win blowing, the current primitive infrastructure cannot detect the circumstance. Besides, the building operators have to visit every single room to have information about the state or control the room facilities instead of having any control interface for a whole building.

System requirements
In general, the prototyped is expected to satisfy the pre-defined requirements to modernize the conventional lecture rooms as follow:

- Providing Automation and Smart features.
- Offering Management features.
- Serving Users, Building managers and Developers.
- Recruiting modern technologies and powerful platforms.
- Scalability, Security.

System architecture
In the Fig. 4, there are three major types of entities: Building Server, Room Controllers and the Room Nodes, and connections among entities. As can be seen, the Building Server connects and manages multiple Room Controller instances. In turn, each of the Room Controller can handle multiple instances of Room Nodes. Each of the Node in a room host either actuators or sensors or even both of the two (as controlled objects). The architecture clearly show the scalability of the system to variation in settings of rooms and buildings. The added information here is the technology to enable the connections between the entities such as: Websocket and ZigBee.

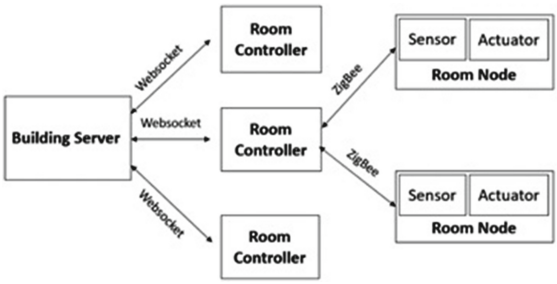


Fig. 4. Overall system architecture for the Smart Lecture Room.

Subsystems distribution
The system can be categorized in three major types of devices: Building Server, Room Controller and Room Nodes. By limiting the development to focus on a lecture room, we only need to design one Room Controller to handle a room and connect it to a Building Server. Even so, the lecture room requires a certain number of Nodes to offer smart functionalities. The proposed Room Nodes are: Main Light Switch Node, Dimmer Node, Sensor Node, Door Lock Node, and Window Node (Fig. 5).

2.3 Design Phase

Room Node general architecture
All the Room Nodes in the system share a common hardware platform in Fig. 6. The variations depend on the peripherals that are attached to. As can be seen in the figure, the chosen Arduino Mega and the ChibiOS play the most important role in each of the subsystem. Besides, a node has XBee shield and its module to perform communication to the outside world. To interact with the controlled object, the sensor and actuator are the solution. A Room Node can have the presence of both sensor and actuator or one of the two.

In the Fig. 7, the implementation of software architecture solution for the MCU nodes in the architecture is proposed. This recalls the employment of Real Time Operating System. As selected, the ChibiOS is used to realize this architecture. By

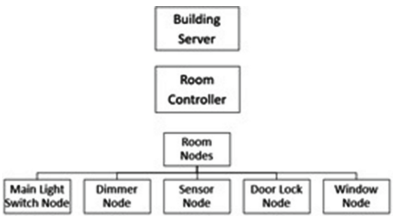


Fig. 5. There are 7 subsystems in total, after the subsystems distribution.

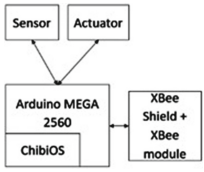


Fig. 6. General hardware architecture for the Room Nodes.

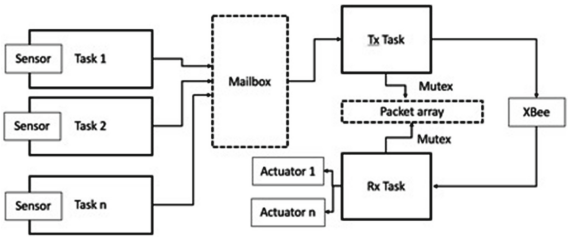


Fig. 7. Generic Multi-tasking, two ways, enhanced reliability communication for ChibiOS on Arduino MEGA platform.

applying the architecture, a Room Node can operate as it expected functions (handle sensors, actuators, and XBee communication module) with added features of two ways communication, multitasking environment, and enhanced transmission reliability. All the Room Nodes in the system use the same multitasking software architecture with modification to adapt to its services.

Room Controller and Building Server

In terms of hardware platform, the Room Controller is constructed up on the Raspberry Pi credit card sized computer (Fig. 8). The Raspberry Pi provides HDMI output to the touch display. The touch action of the user on the display is converted to mouse action by a converter and fed to the Pi using USB interface. The GPIO pins of the Pi are used to connect to the buzzer peripheral. The Pi also have a Websocket to the Building Server via Ethernet interface and XBee module with a FTDI adapter via remained USB port.

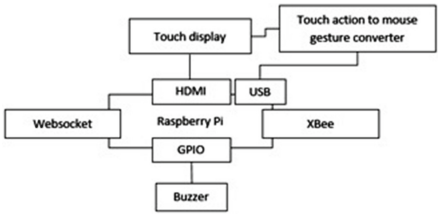


Fig. 8. The hardware architecture of the Room Controller.

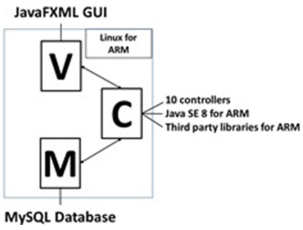


Fig. 9. The software architecture of the Room Controller.

The software architecture of the Room Controller (RC) plays utmost important role in the system (Fig. 9). The software of the RC must be able to handle complicated communications and control strategies to operate the room itself. To do so, the MVC (Model View Controller) architecture is employed. The selected solution for View is JavaFXML, and for Model is the MySQL Server database. The heart of the system are Controllers.

3 Results

As can be seen from Fig. 10, the Room Controller has a display to provider touch HD interface to user. On top of it is the connection for the XBee module to communicate with the Room Nodes. Next to it is the power plug for the Mail Light Switch and Dimmer Node because those devices need the high voltage power source. Under the power plug is the USB hub to power low voltage for operations of the MCU in each of the Nodes. The hub also powers the display and the Pi. In this set up all the power of the Nodes and display for RC come from one source but in practice they must be separated. On the right hand side, there are the Main Switch Node, Dimmer Node, Windows Node and the Door Lock Node. On the top of the panel is the light bulbs to simulate the control of the Main Switch and the Dimmer. On the far right, there is the windows and main door to simulate the contact magnet open close state. At the right bottom are RFID keys to simulate user access.

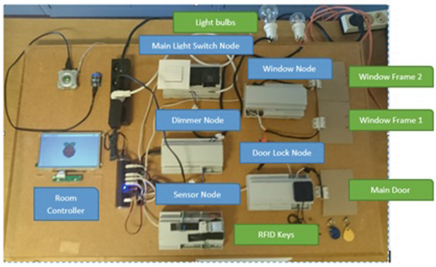


Fig. 10. Overall Smart Lecture Room system.

On the other side of the panel are the other part of the RC with the Pi underneath the display, power cord and the Ethernet cable to connect the room with the Building Server.

The Smart Lecture Room offers the functionalities (Fig. 11):

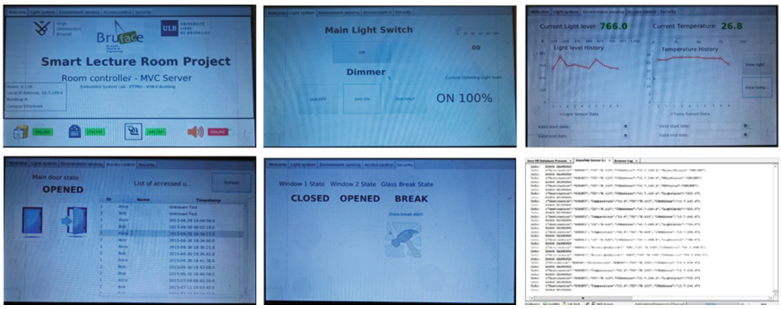


Fig. 11. From Left-To-Right: Welcome tab of the SLR to provide basic information, Light system tab of the SLR to offer controls on lighting system, Environment sensing tab of the SLR to offer overview of environmental information, Access Control tab of the SLR to show current state of main door and entered users, Security tab of the SLR to offer current states of two windows frame and its glass state, The received JSON packets in back end Building Server

- One central control interface allows users to manage the room.
- Toggling the main light switch from GUI.
- The manual switching state action on the electrical light switch from users are reflected in the light control system.
- Switching on, off, and dimming the lights in other group.
- The environmental information including current light level and temperature is displayed in real-time.
- The historical measured information can be viewed in graphs.
- The state of the main door is displayed on the GUI together with a list of accessed user to the room with time-stamp.
- The user with his/her RFID, which data is stored in database, can open the door within 10 s.
- The states of the two window frames are reported to the Room Controller to display on GUI at right moment of state changes.
- The glass of the windows, if in any case of broken, can be detected and alerted.
- Providing Automation and Smart features.
- Database logging on accessed users, sensing data, control actions and so on.
- The room can automatically perform the saving of energy on controlled light system with time-based strategy.
- The motion is detected for the energy saving smart features of the room.
- The Smart Lecture Room provides wired Websocket connection to update a full image and partial images of the room as services (Service Oriented Architecture).

4 Performance Evaluations

4.1 Analysis of Room Nodes

The static behavior of the software variables can be measured by the calculation from the Arduino IDE to achieve the Flash memory usage graph (Fig. 12). As can be seen the control logic in each of the Nodes consumes less than 10 % of the total memory in term of static variables. However, the important information about the dynamical behavior of the memory such as stack allocation when the tasks switch the context or the allocation of local variables when the functions is called and returned is not trivial. The information requires a high level of analysis and out of the scope of this analysis.

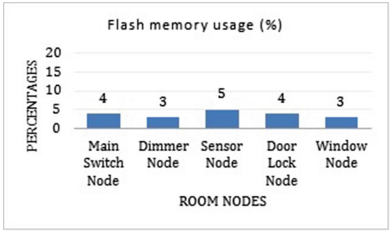


Fig. 12. The usage of Flash memory of the Arduino MEGA 2560 in Room Nodes.

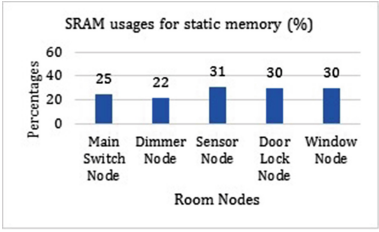


Fig. 13. The SRAM consumption of the Nodes with their multitasking code architecture.

In total 8 k bytes of the SRAM, the usage of the static consumption in the memory is indicated for the Nodes as in the graph (Fig. 13). The local variables and stack usages have dynamical behavior and hard to be measured offline. To have a picture of the memory usage in run time, or in steady state requires more professional tools. From the graph, we can see that the maximum usage or the static memory is less than 50 %, with more intensive tool in analysis on the fly memory consumption in run time can propose a possibility of selecting other MCU platforms with lower unused space.

4.2 Analysis of Room Controller

In term of loaded classes, those classes including the native Java libraries, third party added ones for the application and the classes from the application itself (Fig. 14). That is the reason why, the number of the classes ramp up to over 4000 instances after few minutes from start up. At the steady state, the device uses 4457 classes and 57 threads. The threads will increase when user takes actions on the GUI.

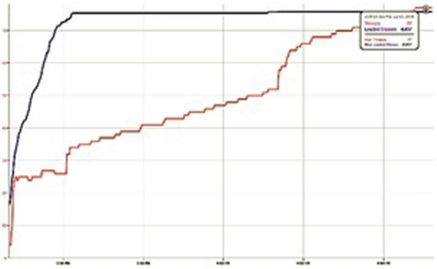


Fig. 14. The Raspberry memory performance at steady state.

4.3 System Cost Evaluation

As can be seen in Fig. 15, the Room Controller cost accounted almost for 30 % of the total cost due to its components. The other nodes are quite equal in costs. However, the difference between the RC and a Room Nodes is quite slightly. A better system should have much large the differences between the Room Controller and its Nodes. This is obvious when a room may contain more Room Nodes and even more Nodes with identical nodes to apply on different facilitates. For instance, a room has 4 different radiators, or two main doors. If the cost of any single node is too expensive, the scalability of the room may have a negative effect on the total bill.

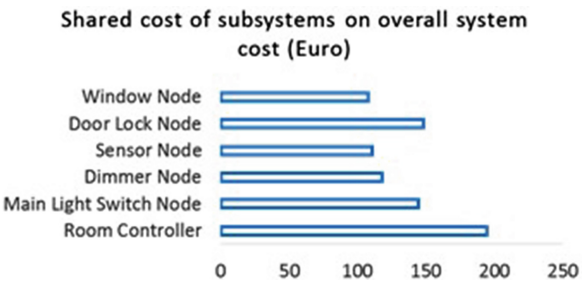


Fig. 15. The costs of subsystems compare to total system cost.

5 Conclusions

A complete solution for a Smart Lecture Room has realized and satisfied the requirements at early stage, which has developed from the abstract information to the real devices. The pre-prototyping conceptual design has been an early proper guidance for the realization of the system. The post-prototyping analysis has made room for improvements for the system.

References

1. Camacho, R., Carreira, P., Lynce, I., Resendes, S.: An ontology-based approach to conflict resolution in home and building automation systems. Expert Syst. Appl. **41**, 6161–6173 (2014)
2. Kastner, W., Neugschwandtner, G., Soucek, S., Michael Newman, H.: Communication systems for building automation and control. Proc. IEEE **93**, 1178–1203 (2005)
3. Marsa-Maestre, I., Lopez-Carmona, M.A., Velasco, J.R., Navarro, A.: Mobile agents for service personalization in smart environments. J. Netw. **3**(5), 30–41 (2008). p. 04
4. Kim, W.H., Lee, S., Hwang, J.: Real-time energy monitoring and controlling system based on ZigBee sensor networks. Procedia Comput. Sci. **5**, 794–797 (2011). International Symposium on Intelligent Systems Techniques for Ad hoc and Wireless Sensor Networks (IST-AWSN)
5. James, F., Smit, H.: Service-oriented architectures for devices. In: The SIRENA View, 3rd IEEE International Conference on Industrial Informatics (INDIN) (2005)
6. Ngo, L.: Service-oriented architecture for home networks. Helsinki University of Technology (2007)
7. Teich, T., Wolf, S., Neumann, T., Junghans, S., Franke, S.: Concept for a service-oriented architecture in building automation systems. In: 24th DAAAM International Symposium on Intelligent Manufacturing and Automation (2013)
8. Ingle, A.J., Gawali, B.W.: Status of wireless technologies used for designing home automation system – a review. Int. J. Adv. Comput. Sci. Appl. (IJACSA) **2**(7), 142–146 (2011)
9. Al-Qutayri, M.A., Jeedella, J.S.: Integrated wireless technologies for smart homes applications. Khalifa University of Science, Technology, and Research United Arab Emirates
10. Anwaarullah, S., Altaf, S.V.: RTOS based home automation system using Android. Int. J. Adv. Trends Comput. Sci. Eng. **2**(1), 480–484 (2013). Special Issue of ICACSE 2013 - Held on 7–8 January, 2013 in Lords Institute of Engineering and Technology, Hyderabad (2013)
11. Thomas, C.J.: Intelligent sensor based building automation and energy management. Int. J. Adv. Res. Comput. Sci. Softw. Eng. **3**(8), 848–851 (2013)
12. Song, G.: Design of a networked monitoring system for home automation. IEEE Trans. Consum. Electron. **53**, 933–937 (2007)
13. Gill, K., Yang, S.-H., Yao, F., Xin, L.: A ZigBee-based home automation system. IEEE Trans. Consum. Electron. **55**(2), 422–430 (2009)

Solving Navier-Stokes Equation Using FPGA Cellular Neural Network Chip

Duc-Thai Vu(✉), Le Hung Linh, and Nguyen Mai Linh

University of Information and Communication Technology-Thai Nguyen University, Thai Nguyen, Vietnam
{vdthai,lhlinh,nmlinh}@ictu.edu.vn

Abstract. This paper presents a method of using Cellular Neural Network (CNN) for solving hydraulic Navier-Stokes equations, which are a set of three partial differential equations with 3 functional variables and each function has three variables for time and space. The paper has 5 parts: part 1 is the introduction; Part 2 discusses the Navier-Stokes for flow through narrow slots; Part 3 analyzes and finds templates for designing CNN architecture; Part 4 provides some solutions for optimizing resources and speeding up computing; Part 5 sets up CNN chip for implementing and computing results; the last part is conclusion and developing trends.

Keywords: Cellular Neural Network · Navier-Stokes equation · Water flow · FPGA · Partial differential equation

1 Introduction

In the context of global warming, thunderstorms cause flash flood and flooding, which leads to serous losses in terms of human life and properties. Calculations of flow phenomena play an important role in forecasting and building disaster prevention systems. Navier -Stokes are one set of the mathematical equations describing hydraulic flow phenomena and have been developed for a long time and varied in accordance with each practical application. These equations have been solved by many researchers manually and computationally. However, this paper deals with Navier-Stokes equations describing water flows which changes velocity and directions when passing through a narrow slot such as a flood discharge outlet [3,10].

Cellular Neural Network is compact on-chip architecture capable of parallel calculation of large-scale physical processes. The CNN chip has learning capability to improve computational performance like other artificial neural networks [8,9]. CNN chip can be made by configuring on programmable FPGA chip. There have been applications of CNN technology in many fields published at International Workshop on Cellular Nanoscale Networks and their Applications (CNNA) every two years. Solving partial differential equations on CNN technology has been widely done over the world. In Vietnam, there have been some

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_60

studies on making CNN chip using FPGA technology to successfully solve partial differential equations Saint Venant [4,5]. It is feasible and reliable to create CNN chip for solving Navier-Stokes equation. Using VHDL language to configure FPGA to make CNN chip is big advantages for creating prototypes at the beginning.

2 Navier-Stokes Equations

2.1 Physico-Mathematical Model of Navier-Stokes Equations

In hydraulics, many flow models have been studied such as flows in canals, rivers, and streams; flows at river mouths; flows in an area of sea, ocean... in order to control the flow for preventing disasters, saving water and exploiting energy of the flow. Mathematical models of those phenomena consist of differential equations and partial differential equations like Saint venant equations or Navier-Stokes equations [3].

Navier-Stokes equations in this study are 3 partial differential equations of functional variables representing water height, and flow velocity in x-and y-directions. The emperical model is a flow through a small conduit outlet, which then diffuses in two directions Ox and Oy [10].

Solving Navier-Stokes equations on electronic devices requires the discretion of continuity model by difference method. To have the accuracy, difference intervals must be small enough. However, too small difference intervals increase the calculation complexity and time. If the physical parallel calculation on CNN is used, the above difficulties will be overcome.

2.2 Desciption Equations in Navier-Stokes Equations

a. Equations Describing Height of Water

$$\frac{\partial \rho z_w}{\partial t} + \frac{\partial \rho q_x}{\partial x} + \frac{\partial \rho q_y}{\partial y} = \rho q_A \tag{1}$$

Assume that the height of water is taken from the bottom of the flow which is regarded as the origin of the coordinate system, so z_w has no negative values.

b. Mometum Equations in x-Direction

$$\frac{\partial \rho q_x}{\partial t} + \frac{\partial}{\partial x}\left(\rho\beta\frac{q_x^2}{d}\right) + \frac{\partial}{\partial y}\left(\rho\beta\frac{q_x q_y}{d}\right) + \rho g d\frac{\partial z_w}{\partial x} + \rho g d S_{fx} - \tau_{wx} - \frac{\partial}{\partial x}\left(\rho K_L \frac{\partial q_x}{\partial x}\right) - \frac{\partial}{\partial y}\left(\rho K_T \frac{\partial q_x}{\partial y}\right) = 0 \tag{2}$$

c. Momemtum Equations in y-Direction

$$\frac{\partial \rho q_y}{\partial t} + \frac{\partial}{\partial y}\left(\rho\beta\frac{q_y^2}{d}\right) + \frac{\partial}{\partial x}\left(\rho\beta\frac{q_x q_y}{d}\right) + \rho g d\frac{\partial z_w}{\partial y} + \rho g d S_{fy} - \tau_{wy} - \frac{\partial}{\partial y}\left(\rho K_L \frac{\partial q_y}{\partial y}\right) - \frac{\partial}{\partial x}\left(\rho K_T \frac{\partial q_y}{\partial x}\right) = 0 \tag{3}$$

2.3 Analysing and Designing CNN to Solve the Equations

To simplify the variables in Navier-Stokes equations, let: Height of water be $h = z_w$; velocity in $x-$axis be $u = q_x$, in $y-$axis be $v = q_y$. Assume that the amount of water losses because of evaporation and penetration is trivial or $q_A = 0$; kinetic influence of turbulent values between velocity in the direction from Oy to Ox (or Ox to Oy) is trivial since horizontal velocity is small enough to be ignored.

$$\frac{\partial h}{\partial t} + \frac{\partial u}{\partial x} + \frac{\partial v}{\partial y} = 0 \Leftrightarrow \frac{\partial h}{\partial t} = -\frac{\partial u}{\partial x} - \frac{\partial v}{\partial y} \tag{4}$$

$$\frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(\beta\frac{u^2}{d}\right) + \frac{\partial}{\partial y}\left(\beta\frac{uv}{d}\right) + gd\frac{\partial h}{\partial x} + gdS_{fx} - \frac{\tau_{wx}}{\rho} - \frac{\partial}{\partial x}\left(K_L\frac{\partial u}{\partial x}\right) \tag{5}$$

$$\Leftrightarrow \frac{\partial u}{\partial t} = \frac{\partial}{\partial x}\left(K_L\frac{\partial u}{\partial x}\right) - \frac{\partial}{\partial x}\left(\beta\frac{u^2}{d}\right) - \frac{\partial}{\partial y}\left(\beta\frac{uv}{d}\right) - gd\frac{\partial h}{\partial x} + \left(\frac{\tau_{wx}}{\rho} - gdS_{fx}\right)$$

$$\frac{\partial v}{\partial t} + \frac{\partial}{\partial y}\left(\beta\frac{v^2}{d}\right) + \frac{\partial}{\partial x}\left(\beta\frac{vu}{d}\right) + gd\frac{\partial h}{\partial y} + gdS_{fy} - \frac{\tau_{wy}}{\rho} - \frac{\partial}{\partial y}\left(K_L\frac{\partial v}{\partial y}\right) = 0 \tag{6}$$

$$\Leftrightarrow \frac{\partial v}{\partial t} = \frac{\partial}{\partial y}\left(K_L\frac{\partial v}{\partial y}\right) - \frac{\partial}{\partial y}\left(\beta\frac{v^2}{d}\right) - \frac{\partial}{\partial x}\left(\beta\frac{vu}{d}\right) - gd\frac{\partial h}{\partial y} + \left(\frac{\tau_{wy}}{\rho} - gdS_{fy}\right)$$

a. Differencing Equations Following Taylor Formular

Spatial finite differences use finite difference grid with difference interval in $x-$axis as Δx; in $y-$axis as Δy and apply Taylor difference formulas for Eqs. (4), (5), (6), we have difference equations corresponding the equations:

$$\frac{\partial h_{ij}}{\partial t} = \frac{u_{i+1,j} - u_{i-1,j}}{2\Delta x} - \frac{v_{i,j+1} - v_{i,j-1}}{2\Delta y} \tag{7}$$

$$\frac{\partial u_{i,j}}{\partial t} = -\frac{\beta}{d}\left[\frac{u_{i+1,j}}{2\Delta x}u_{i+1,j} - \frac{u_{i-1,j}}{2\Delta x}u_{i-1,j}\right] - \frac{\beta}{d}\left[\frac{v_{i,j+1}}{2\Delta y}u_{i+1,j} - \frac{v_{i,j-1}}{2\Delta y}u_{i-1,j}\right] - gd\frac{h_{i+1,j} - h_{i-1,j}}{2\Delta x} - gdS_{fx} + \frac{1}{\rho}\tau_{wx}K_L\frac{u_{i+1,j} - 2u_{i,j} + u_{i-1,j}}{\Delta x^2} \tag{8}$$

$$\frac{\partial v_{i,j}}{\partial t} = -\frac{\beta}{d}\left[\frac{v_{i,j+1}}{2\Delta y}v_{i,j+1} - \frac{v_{i,j-1}}{2\Delta y}v_{i,j-1}\right] - \frac{\beta}{d}\left[\frac{u_{i+1,j}}{2\Delta x}v_{i,j+1} - \frac{u_{i-1,j}}{2\Delta x}v_{i,j-1}\right] - gd\frac{h_{i,j+1} - h_{i,j-1}}{2\Delta x} - gdS_{fy} + \frac{1}{\rho}\tau_{wy}K_L\frac{v_{i,j+1} - 2v_{i,j} + v_{i,j-1}}{\Delta y^2} \tag{9}$$

b. Designing a Sample of CNN

Based on CNN state equations and difference equations [8,9], we can have CNN templates for layers h, u, v :

+ Layer h :

$$A^{hu} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{1}{2\Delta x} & 0 & \frac{-1}{2\Delta x} \\ 0 & 0 & 0 \end{bmatrix} \quad A^{hv} = \begin{bmatrix} 0 & \frac{1}{2\Delta y} & 0 \\ 0 & 0 & 0 \\ 0 & \frac{-1}{2\Delta y} & 0 \end{bmatrix} \quad B^h = 0;\ z^h = 0 \tag{10}$$

+ Layer u :

$$A^{uv} = \begin{bmatrix} 0 & \frac{\beta u_{i,j-1}}{2d\Delta y} & 0 \\ 0 & 0 & 0 \\ 0 & \frac{-\beta u_{i,j+1}}{2d\Delta y} & 0 \end{bmatrix} \quad A^{uh} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{gd}{2\Delta x} & 0 & \frac{-gd}{2\Delta x} \\ 0 & 0 & 0 \end{bmatrix} \quad B^u = \begin{bmatrix} 0\,0\,0 \\ 0\,1\,0 \\ 0\,0\,0 \end{bmatrix} \quad z^u = 0$$

$$A^u = \begin{bmatrix} 0 & 0 & 0 \\ \frac{\beta u_{i-1,j}}{2d\Delta x} + \frac{K_L}{\Delta x^2} & gd\frac{n^2(u_{ij}^2 + v_{ij}^2)^{\frac{1}{2}}}{d^{\frac{1}{3}}} + \frac{1}{R^u} + \frac{4K_L}{\Delta x^2} & \frac{-\beta u_{i+1,j}}{2d\Delta x} + \frac{-K_L}{\Delta x^2} \\ 0 & 0 & 0 \end{bmatrix} \tag{11}$$

+ Layer v :

$$A^{vh} = \begin{bmatrix} 0 & \frac{gd}{2\Delta y} & 0 \\ 0 & 0 & 0 \\ 0 & \frac{gd}{2\Delta y} & 0 \end{bmatrix} \quad A^{vu} = \begin{bmatrix} 0 & 0 & 0 \\ \frac{\beta u_{i-1,j}}{2d\Delta x} & 0 & \frac{-\beta u_{i-1,j}}{2d\Delta x} \\ 0 & 0 & 0 \end{bmatrix} \quad B^v = \frac{1}{\rho}\tau_{wy} \begin{bmatrix} 0\,0\,0 \\ 0\,1\,0 \\ 0\,0\,0 \end{bmatrix} \quad z^v = 0$$

$$A^v = \begin{bmatrix} 0 & \frac{\beta v_{i,j+1}}{2d\Delta y} + \frac{K_L}{\Delta y^2} & 0 \\ \frac{K_L}{\Delta y^2} & gd\frac{n^2(u_{i,j}^2 + v_{i,j}^2)^2}{d^{\frac{1}{3}}} + \frac{1}{R^v} + \frac{K_L}{\Delta y^2} & \frac{-K_L}{\Delta y^2} \\ 0 & \frac{-\beta v_{i,j+1}}{2d\Delta y} + \frac{K_L}{\Delta y^2} & 0 \end{bmatrix} \tag{12}$$

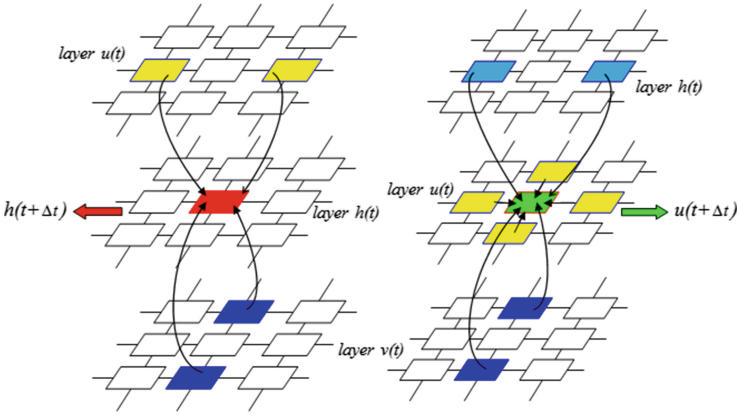


Fig. 1. Arithmetic unit for function h and u

c. Designing Hardware Architecture of CNN to Solve Navier – Stokes Equations

Based on templates found in (10), (11), (12), we can design architecture for CNN. The architecture of neural network used to Navier-Stokes equations is 3-layered CNN2D. From the general design model, we can create an arithmetic unit for each layer and links to perform parallel calculation on the whole chip. Figure 1 shows the architecture of layer h, u (the layer v is similar to u) [2,6,7].

3 Configure CNN Chip from FPGA

The emperical problems which need to be solved is that: Firstly, because resources of FPGA are limited and the space requiring for calculation is large, it is necessary to divide the space into smaller areas for combining sequential and parallel computing; Secondly, the division and combination of boundary areas are required to be continous and not to lead to incorrect results due to coordinating time of computing Thirdly, real-time data exchange which is unsolved as in the previous reports [1,4–6]. The CNN chip proposed in this paper has solved all the previous problems.

3.1 Proposed System Architecture

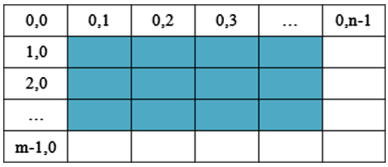
The control unit controls the activities of the whole system set by the algorithm. When the computing space is divided into areas, there are buffer areas (boundary). The buffer area for the whole computing space has size of $M \times N$ as shown in the following matrix:

0,0	0,1	...	0,N-1
1,0			
...			
M-1,0			

The white area is input buffer area, the grey part is the area which requires to be processed by CNN chip. Arithmetic unit of CNN chip has size of $(M-2)\times(N-2)$ cells processing data for the grey area which is inside the input buffer area.

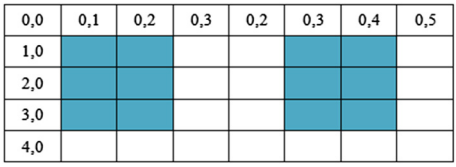


For the purpose of simple installation, we select $M = 3$. N is selected depending on the resources of FPGA chip we use (large or small). Here, assume that $N = 4$, thus, CNN output buffer area has the size of $1 \times N$.
Data which need processing sent from PC have the size of $m \times n$:

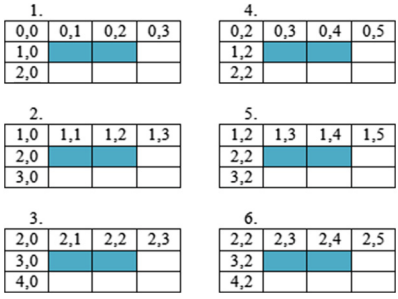


3.2 Separate Computing Space

Assume that $m = 5, n = 6$, the white part is boundary, the grey part is the data area requiring to be processed. The vertical boundaries will be added for storing temporary data.



Temporary vertical boundaryies are added to the data structure similar to CNN input buffer. Temporary horizontal boundaries are not necessary due to data retrieval algorithms of control unit. The data after being added with temporary vertical boundareis will be sent to Memory unit into h, u, v. Then, the control unit will control to read each data of the same size as CNN input buffer from Memory unit into h, u, v (in case that $m \times n = 5 \times 6, M \times N = 3 \times 4$) detailed as follows:



These in turns will be sent to CNN arithmetic unit to be processed in pipeline mode. The processing results will be written in Memory unit h, u, v.

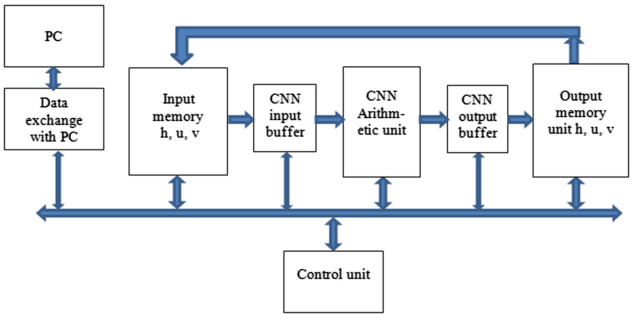


Fig. 2. The architecture of CNN chip

Next, the results after being processed in Memory unit h, u, v will be updated with tempory vertical boundaries and brought back to memory unit h, u, v for the next processing turn (Fig. 2).

4 Installation Results

4.1 General Circuit Diagram

It is known that CNN work in global and local manners. The Control unit controls all input/output manupulations and calculations. Especially, in this case, it is the division and cooperation of calculation space in terms of time and scanning of the whole space. It is noted that interval is very small and measured by clock pulses (Figs. 3, 4 and 5).

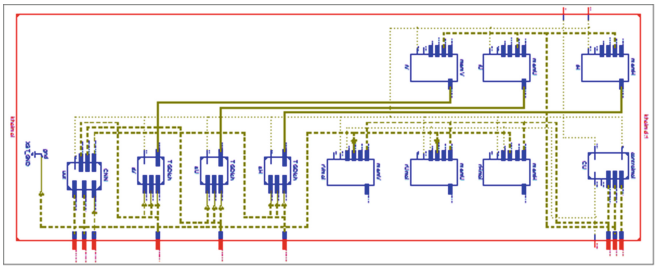


Fig. 3. The block design of CNN chip

4.2 CNN Arithmetic Unit Diagram

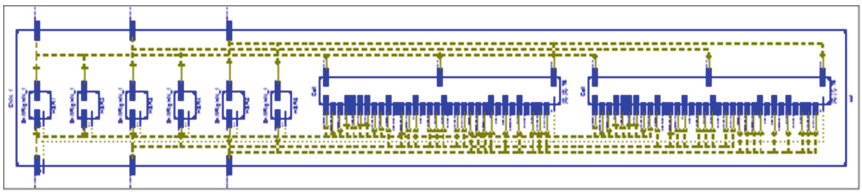


Fig. 4. The CNN Arithmetic unit

4.3 Diagram of One Cell Configured CNN

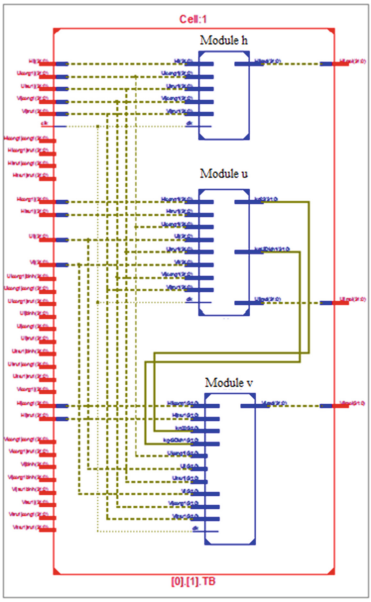


Fig. 5. One computing block for h, u, v

4.4 Computing Simulation

The data sent to chip from port is presented in 32-bit floating real numbers. The bit stream is louded in pipeline mode for set of h, u, v variables. For example, input values of h are shown in Fig. 6 as follow:

The computing results of h are shown in Fig. 7

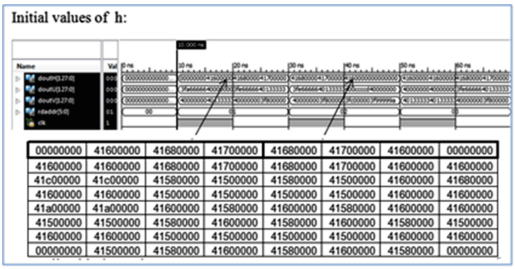


Fig. 6. The computing results in pulse signal

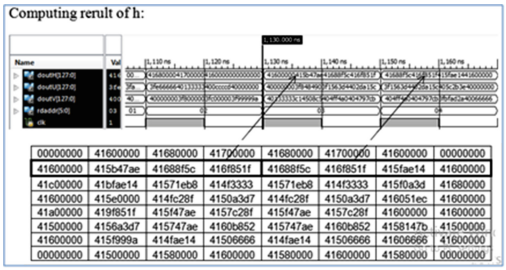


Fig. 7. The computing results in pulse signal

5 Conclusions and Future Trends

This paper introduces a solution to config CNN chip to solve Navier-Stokes equations. It also presents an effective solution to the temporary boundary problem when it is required to divide the big data to be processed into many subdata. The processing of the whole big data is based on the CNN calculation of each subdata. With the input data of 32-bit floating point real number and FPGA chip Vertex 6 XCVL240T-1FFG1156 by Xilinx, the team has successfully installed a CNN of 1×12 cells. The installation results show that the effectiveness of this solution mainly lies on the expansion of calculation space and resources saving. Although the detailed calculation results are not presented due to the limited space of this paper, the research team is sure about the accuracy of the calculation which is published in some previous works by the research team. The model can be further developed to feasibly solve problems on automatic control systems.

References

1. Thai, D.V., Anh, B.T., Duong, V.D.: Develop somes application of Cyclone - DE2C35 chip. J. Sci. Technol. Thai Nguyen Univ. **10–140**, 103–108 (2015)
2. Thai, V.D., Phanthavong, S.: Design template in cellular neuron network. J. Sci. Technol. Thai Nguyen Univ. **116**(2), 3–8 (2014)

3. Kenig, C.E., Koch, G.S.: An alternative approach to regularity for the Navier-Stokes equation in critical spaces. Annales de l'Institut Henri Poincare (C) Non Linear Anal. **28**(2), 159–187 (2010). Elsever Science Direct
4. Thai, V.D., Cat, P.T.: Equivalence and stability of two-layer cellular neural network solving saint venant 1D equation. In: Proceeding (ISI) of 11th International Conference on Control, Automation, Robotics and Vision (ICARCV 2010), Singapore, 7–10 December 2010, pp. 704–709. http://ieeexplore.ieee.org
5. Thai, V.D., Cat, P.T.: Solving two-dimensional saint venant equation by using cellular neural network. In: Proceeding of the 7th Asian Control Conference - ASCC 2009, Hongkong, China, pp. 1258–1263 (2009). http://ieeexplore.ieee.org
6. Kocsárdi, S., Nagy, Z., Csík, A., Szolgay, P.: Two-dimensional compressible flow on emulated digital CNN-UM. In: Proceeding of 11th International Workshop on CNN and Their Applications, pp. 169–174 (2008)
7. Nagy, Z., Szolgay, P.: Emulated digital CNN-UM implementation of a barotropic ocean model. In: Proceeding 2004 IEEE International Joint Conference, vol. 4, pp. 101–108 (2004)
8. Chua, L.O., Yang, L.: Cellular neural networks: theory. IEEE Trans. Circ. Syst. **35**(10), 1257–1272 (1988)
9. Chua, L.O., Yang, L.: Cellular neural networks: application. IEEE Trans. Circ. Syst. **35**, 1273–1290 (1988)
10. Rusin, W.M.: On solution to Navier-Stokes equation in critical spaces. Thesis of Doctor Philosophy (2010). http://conservancy.umn.edu/.../Rusin_umn_0130E_11277.pdf

Special Characters of Vietnamese Sign Language Recognition System Based on Virtual Reality Glove

Diep Nguyen Thi Bich[1(✉)], Trung-Nghia Phung[1], Thang Vu Tat[2], and Lam Phi Tung[2]

[1] Thai Nguyen University of Information and Communication Technology, Thai Nguyen, Vietnam
{ntbdiep,ptnghia}@ictu.edu.vn
[2] Institute of Information Technology, Vietnam Academy of Science and Technology, Hanoi, Vietnam
{vtthang,tunglam}@ioit.ac.vn

Abstract. In this paper, we introduce a method of recognition numbers and special characters of Vietnam sign language. We address a development of a glove-based gesture recognition system. A sensor glove is attached ten flex sensors and one accelerometer. Flex sensors are used for sensing the curvature of fingers and the accelerometer is used in detecting a movement of a hand. Depending on the hand's postures, i.e., vertical, horizontal, and movement, sign language of numbers and special characters can be divided to group 1, 2, and 3, respectively. Firstly, the hand's posture is recognized. Next, if the hand's posture belongs to either group 1 or group 2, a matching algorithm is used to detect a number or one of special characters. If the posture belongs to group 3, a dynamic time warping algorithm is applied. The use of our system in recognizing Vietnamese sign language is illustrated. In addition, experimental results are provided.

Keywords: Recognition · Vietnamese sign language · Number · Special characters · Vituarl reality glove

1 Introduction

There are about 360 millions of deaf people in the world, equivalent to 5 % of the total world population [13]. Most of deaf people are poverty because of restricted educational opportunities and the poor communication. Today, researchers are increasingly paying attention to construction tools translate sign language - the language of the deaf, especially the field of investigation of hand shape and gesture recognition because it is so useful in several applications, e.g., telemanipulation, sign language translation, robotics [9], etc. In this paper, we aim to develop a glove-based gesture recognition system that allows recognizing Vietnamese sign language (VSL), performed by a user with a single hand, using a

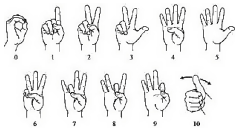
© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_61

data glove as an input device. We focus on the classification and recognition of gestures that represent Number and special characters of Vietnamese sign language. Among the vast variety of existing approaches for hand shape and gesture recognition, methods using sensing gloves have proven to be remarkably successful [1,7]. A survey of glove-based system and their applications is presented in [4]. Mehdi and Khan [8] used a sensor glove to capture the signs of American sign language (ASL) performed by a user and translate them into sentences of English language. In addition, artificial neural networks (ANNs) are used to recognize the sensor values coming from the sensor glove. ANNs have been used for both (static) postural classification [3] and gesture classification [5,12]. A data glove is used for recognition the Japanese alphabets [10], for the Chinese language [2], etc. Vietnamese vocabulary is more complicated than English alphabet system because of more signs for VSL in comparison with ASL. Special characters are only available in Vietnamese. Bui and Nguyen [10,11] created 22 fuzzy rules to classify Vietnamese sign language postures. They used a sensing glove that is attached six accelerometers and a basic stamp microcontroller in recognizing Vietnam number and special characters sign language. In this paper, we aim to develop a glove-based gesture recognition system in which data glove are used in classification and recognition numbers and special characters in Vietnamese sign language. The glove has two main parts, i.e., sensors (flex sensors and an accelerometer) and a system of data processing and communication. Firstly, the hand's posture is detected. Depending on the hand's posture, i.e., vertical, horizontal, and movement, sign language of alphabets are divided into group 1, 2, and 3, respectively. In the next stage, if the hand's posture belongs to either group 1 or 2, a matching algorithm is used to detect a letter. If the posture belongs to group 3, a letter is recognized by using a dynamic time warping algorithm (DTW). The system of data processing and communication (using microchip Atmega32U) handles data from sensors and then transfer results achieved to PC through USB port. Software running on a PC receives data and then displays an animation of a glove's gestures and a letter recognized. The paper is organized as follows: our data set and sensing glove are introduced in Sect. 2. In Sect. 3, our recognition system of Vietnamese sign language is described in detail. Experimental results are presented in Sect. 4. Finally, conclusions are drawn in Sect. 5.

2 The Data Set and Sensing Glove

2.1 The Data Set

Our data set are numbers and specials character in Vietnamese sign language (VSL). The numbers performances in Vietnamese sign language are different to other such as: American (ASL), Chinines (CSL). The expressing numbers in VSL, similar ASL and CSL with number 0 to 5, diffirent with number 6 to 9. Vietnamese alphabet system is more complicated than English alphabet system because more signs are needed for VSL in comparison with ASL.



(a) Numbers in America sign language.



(b) Numbers in Vietnamese sign language.

Fig. 1. Number in sign language.

Some Vietnamese typing tool as Unikey, if you want to type specail character, you must use some letters: w, s, f, r, x, j or use number 1 to 9.

Several specials character in Vietnamese are: acute (’), grave accent (‘), question mark(?), tilde (~). They are only available in Vietnamese.
In this paper, we are going to assess on dataset a list of the following: 0, 1, 2, 3, 4, 5, 6, 7, 8, 9 and acute (’), grave accent (‘), question mark(?), tilde (~).

2.2 A Sensing Glove

The use of the sensor has been presented in [6]. Our sensing glove has two main parts, i.e., sensors (ten flex sensors and one accelerometer) and a system of data processing and communication. There are two flex sensors in one finger. Sensors are fixed in one point then they can move when fingers bent. An accelerometer and a system of data processing and communication (microchip Atmega32U is used) are assembled in one small board that can be immobilized with a user’s wrist. Flex sensors are passive resistive devices that can be used to detect bending or flexing. Flex sensors are analog resisters and work as analog voltage dividers. Inside the flex sensor are carbon resistive elements within a thin flexible substrate. When the substrate is bent, the sensor produces a resistance output relative to the bend radius. An output of a flex sensor is an analog. Ten outputs of flex sensors are connected to ten ADC channels of microchip Atmega32U.

Here, we use an accelerometer, i.e., ADXL345. A function block diagram of ADXL345 is shown in [6]. The ADXL345 is a small, thin, low power, three-axis accelerometer with high resolution (13-bit) measurement up to ± 16 g. The ADXL345 is well suited for mobile applications. It measures the static acceleration of gravity in tilt-sensing application, as well as dynamic acceleration resulting from motion or shock. Digital output data is formatted as 16-bit twos complement and is accessible through either a SPI (3- or 4-wire) or I2C digital interface. Figure 2 depicts X-Y-Z axis of an accelerometer in which the X-axis coincide with the direction of a hand, the Z-axis is taken to be vertical when the hand is in the horizontal plane. An accelerometer returns magnitudes of the projection of vector g to X-Y-Z axis, respectively. These digital output data is accessible through a SPI of Atmega32U.

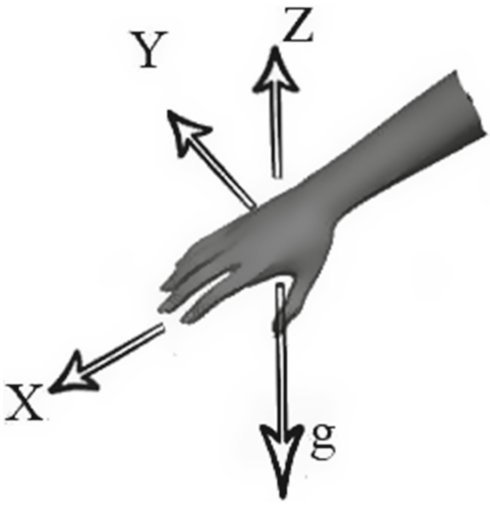


Fig. 2. X-Y-Z axis of an accelerometer in which the X-axis coincide with the direction of a hand, the Z-axis is taken to be vertical when the hand is in the horizontal plane, and g is a gravitational acceleration vector.

3 Recognition of Vietnamese Numbers and Special Characters Sign Language

In this Section, we present our algorithm for classification and recognition of Vietnamese numbers and special characters sign language. The data set that we selected can be divided to three groups depending on the hand's postures: (i) Group 1: when the hand's posture is vertical, which consists of the postures of numbers, i.e., 0, 1, 2, 3, 4, 5, 6, 8 and 9. (ii) Group 2: when the hand's posture is horizontal, i.e., 7. (iii) Group 3: when the hand makes a move, which consists of the postures of letters, i.e., acute (’), grave accent (‘), question mark(?) and tilde ($\tilde{}$). An accelerometer returns values of the projection of a gravitational acceleration vector, g, to 3-axis acceleration sensor. Let (Ax, Ay, Az) be magnitudes of the projection of vector g to X-Y-Z axis, respectively. Let S be a vector of 13 measurement parameters from sensors attached on the glove and is denoted by:

$$\mathbf{S} = [f_{11}\ f_{12}\ f_{21}\ f_{22}\ f_{31}\ f_{32}\ f_{41}\ f_{42}\ f_{51}\ f_{52}\ A_x\ A_y\ A_z]^T$$

where $i = \overline{1,5}$ are values measured from two flex sensors attached on finger i, starting from a thumb to a little finger. Based on signals from sensors attached on the glove, our system recognizes Vietnamese alphabet sign language by a user with a single hand, using the data glove as an input device. Here, flex sensors are used for sensing the curvature of fingers and the accelerometer is used in recognizing the movement of a hand. Firstly, the hand's postures are divided into three groups. Next, if the posture belongs to either group 1 or group 2, the matching algorithm is used to detect a letter. Given a sampling measurement

vector, we calculate a list of errors between the sampling measurement vector and a template vector of each letter belonging to group 1 (or group 2). An output is a letter corresponding to a letter that has the smallest error in the list. If the posture belongs to group 3, the DTW is applied to detect a letter. DTW is an algorithm for measuring similarity between two temporal sequences which may vary in time or speed. Here, DTW is used to find an optimal alignment between the sequences of movement of the hand and the sequences of template movement of sign language of letters under certain restrictions. Our algorithm scheme for classification and recognition is presented in Fig. 5.

3.1 Classification

Assuming that we have n sampling measurement vectors that are recorded continuously from time t0 to tn, Tt, t = [t0,t1,. . . ,tn]. The variance of Ax is determined as follows:

$$\mathrm{Var}(A_x) = \frac{1}{n}\sum_{h=t}^{t+n}\left(A_x^h - \bar{A}_x\right)^2 \tag{B.1}$$

where $\bar{A}_x$ is the expected value, i.e.,

$$\bar{A}_x = \frac{1}{n}\sum_{h=t}^{t+n} A_x^h \tag{B.2}$$

If the variance of Ax, is large than constant ε, the hand is movable. If the variance of Ax is smaller than ε, the hand is immobile and the hand's posture is determined as follows:

$$Hand's\ posture = \begin{bmatrix} Horizontal\ if\ A_x \in (-60,\ 0] \\ Vertical \quad if\ A_x \in [-137,\ -100] \\ NULL \quad Otherwise \end{bmatrix} \tag{B.3}$$

In this paper, $n = 8, \varepsilon = 3$. Figure 4 presents an example of the hand's postures depending on the values of Ax (Fig. 3).

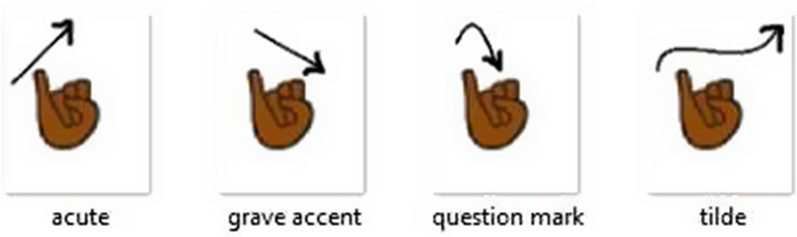


Fig. 3. An example of the hand's gestures corresponding to special charactes of the Vietnamese sign language.

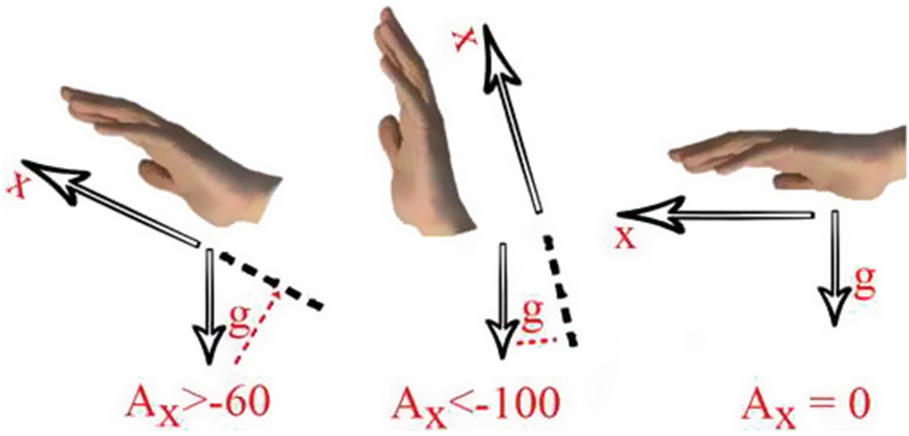


Fig. 4. The hand's postures depending on the values of Ax.

3.2 Recognition

+ If the hand is immobile, we use the template matching method to detect a letter for both cases: the hand's posture is vertical or horizontal. Here, we do not use parameters of an accelerometer because it is used for the classification stage. Let $\mathbf{T}^k = [f_{11}^k\ f_{12}^k\ f_{21}^k\ f_{22}^k\ f_{31}^k\ f_{32}^k\ f_{41}^k\ f_{42}^k\ f_{51}^k\ f_{52}^k\ 0\ 0\ 0]^T$ be a template vector of letter k-th in group 1, where is the number of letters in group 1 $f_{ij}^k,\ i \in [1,5], j \in [1,2]$ is the value measured from a flex sensor. Let be a sampling measurement vector at time t and is denoted by $\mathbf{S}^t$ be a sampling measurement vector at time t and is denoted by

$$\mathbf{S}^t = [f_{11}^t\ f_{12}^t\ f_{21}^t\ f_{22}^t\ f_{31}^t\ f_{32}^t\ f_{41}^t\ f_{42}^t\ f_{51}^t\ f_{52}^t\ A_x^t\ A_y^t\ A_z^t]^T \tag{B.4}$$

Let $\Delta_{t,k}$ be the error of $\mathbf{S}^t$ and $\mathbf{T}^k$ and is calculated as follows:

$$\Delta_{t,k} = \frac{\sum_{i\in[1,5], j\in[1,2]} \left(f_{ij}^t - f_{ij}^k\right)}{10} \tag{B.5}$$

$\arg \min_{k\in[1,\,N_{C1}]} \Delta_{t,k}$ is calculated and then return letter k-th in group 1. The recognition of letters in group 2 is performed similarly. If the hand is movable, the DTW is applied to recognize a letter. Let $\hat{\mathbf{S}}^n = (S^0, ..., S^n)$ be a set of n sampling measurement vectors from time t_0 to t_n, where is a measurement vector at time $t \in (t_0, t_n)$

$$\mathbf{S}^t = [f_{11}^t\ f_{12}^t\ f_{21}^t\ f_{22}^t\ f_{31}^t\ f_{32}^t\ f_{41}^t\ f_{42}^t\ f_{51}^t\ f_{52}^t\ A_x^t\ A_y^t\ A_z^t]^T \tag{B.6}$$

Let $\hat{\mathbf{T}}^{k,m} = (T^{k,t0}, .., T^{k,tm})$, $k = 1, ..., N_{C3}$, be a set of m template vectors from time t_0 to t_n, where is a template vector at time $t \in (t_0, t_n)$ of letter k-th in group 3, where N_{C3} is the number of letters in group 3.

$$\mathbf{T}^{k,t} = [f_{11}^{k,t}\ f_{12}^{k,t}\ f_{21}^{k,t}\ f_{22}^{k,t}\ f_{31}^{k,t}\ f_{32}^{k,t}\ f_{41}^{k,t}\ f_{42}^{k,t}\ f_{51}^{k,t}\ f_{52}^{k,t}\ A_x^{k,t}\ A_y^{k,t}\ A_z^{k,t}]^T \tag{B.7}$$

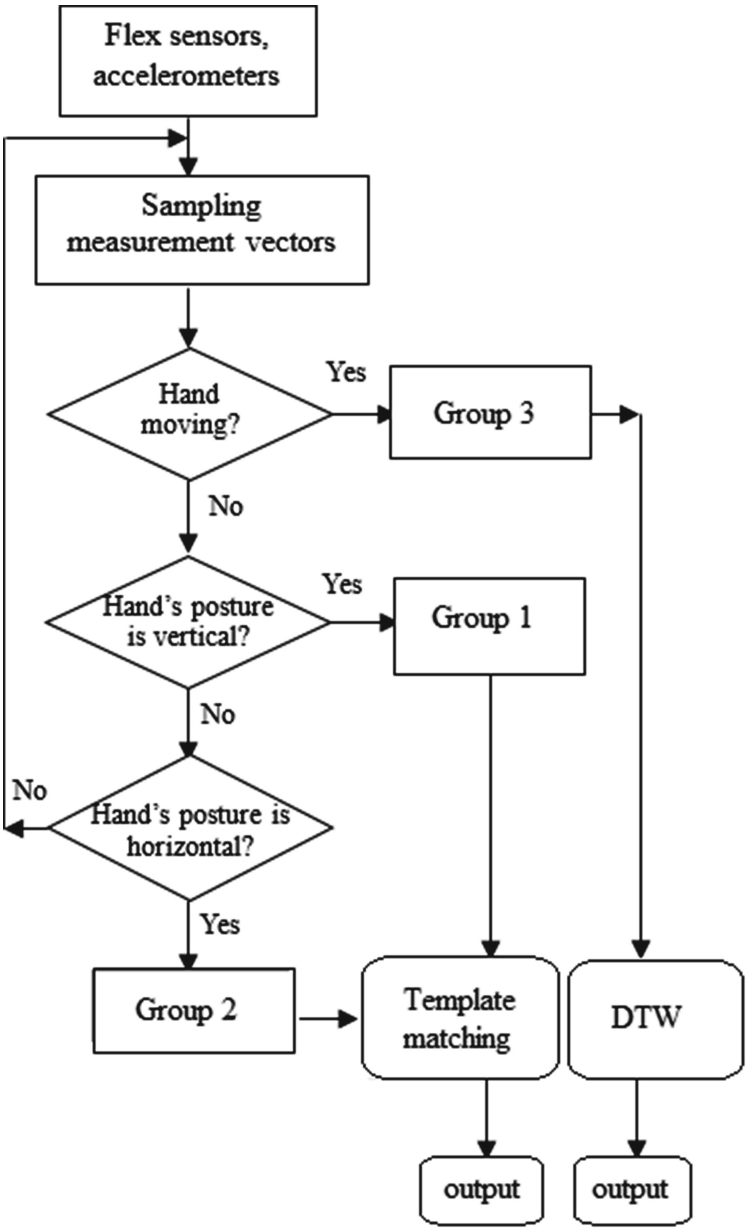


Fig. 5. An algorithm scheme for classification and recognition of Vietnamese numbers and special characters sign language.

Let $\Delta(x, y)$ be the error of $\mathbf{S}^x$ and $\mathbf{T}^{k,y}$, $x \in (t_0, t_n)$, $y \in (t_0, t_m)$ and is calculated as follows:

$$\Delta(x, y) = \frac{\sum_{i\in[1,5], j\in[1,2]} \left(f_{ij}^{x} - f_{ij}^{k,y}\right)}{10} \tag{B.8}$$

Without lost of generality, assuming that t = 1, we have $x = \overline{1,n}$ and $y = \overline{1,m}$. Time-normalized distance is determined as follows:

$$D(\hat{\mathbf{S}}^n, \hat{\mathbf{T}}^{k,m}) = \frac{g(n,m)}{n+m} \tag{B.9}$$

where g(n,m) is calculated recursively as follows:

$$\begin{aligned} g(1,1) &= \Delta(1,1) \\ g(x,1) &= g(x-1,1) + \Delta(x,1) \\ g(1,y) &= g(1,y-1) + \Delta(1,y) \\ g(x,y) &= \min \begin{pmatrix} g(x,y-1) + \Delta(x,y) \\ g(x-1,y) + \Delta(x,y) \\ g(x-1,y-1) + \Delta(x,y) \end{pmatrix} \end{aligned} \tag{B.10}$$

Finally, $\arg \min_{k \in [1,\, N_{C3}]} D(\hat{\mathbf{S}}^n, \hat{\mathbf{T}}^{k,m})$ is calculated and then return letter k-th in group 3.

4 Experimental Results

In this Section, the use of our system in recognizing Vietnamese numbers and special characters sign language is illustrated. We developed a soft-ware running

Table 1. Precision rates of sign language recognition for numbers and special characters of Vietnam sign language

Character	Testing number	Precision number	Precision rate (%)
1	50	48	96
2	50	50	100
3	50	50	100
4	50	50	100
5	50	50	100
6	50	43	86
7	50	50	100
8	50	45	90
9	50	47	94
0	50	46	92
grave accent	50	30	60
acute	50	34	68
question mark	50	35	70
tilde	50	32	64

on a PC in which an animation of the sensing glove and a character detected are shown. Several samples are tested for each letter of the Vietnamese alphabet. Precision rates of sign language recognition for letters are shown in Table 1. The testing process includes steps:
Step 1: We had sign language expert that wear Virtual Reality Glove. Her hand movements under the sign language on our data set.
Step 2: Our group monitoring process on step 1. Based on that we get 50 data types for each symbol is labeled. Data for samples run through the algorithm to obtain the labels.
Step 3: Calculated% of the results obtained, coinciding with the label is correct, the difference with the wrong label available. Thus producing the results in Table 1.

Several characters are recognized with precision rate 100 %, i.e., 2, 3, 4, 5, 7. Four characters, i.e., acute (’), grave accent (‘), question mark(?), tilde (̃) in category 3, have low precision rates because the hand is rotated around Z-axis.

5 Conclusion

In this paper, we focus on recognition numbers and special characters in Vietnamese sign language. We design our system using a data glove that is attached ten flex sensors and one accelerometer. The recognition process has two stages, i.e., recognizing the hand’s posture and detecting numbers and special characters, respectively. Depending on the hand’s posture, either the matching algorithm or the DTW is used to detect a letter. The utility of our system in recognizing Vietnamese sign language is demonstrated. Precision rates of sign language recognition are reported. In future works, we aim to extend our glove-based gesture recognition system for complicated vocabulary in Vietnamese. In the future, we plan to develop the identification system is a large set of signs commonly used in Vietnam sign language. Thereby creating a complete system for the deaf aid.

Acknowledgments. This study was supported by the Ministry of Education and Training of Vietnam (project B2016-TNA-27).

References

1. Parton, B.: Sign language recognition and translation: a multidiscipline approach from the field of artificial intelligence. J. Deaf Stud. Deaf Educ. **11**(11), 94–101 (2006)
2. Xu, D., Yao, W., Zhang, Y.: Hand gesture interaction for virtual training of SPG. In: Proceedings of the International Conference on Artificial Reality and Telexistence, pp. 672–676 (2006)
3. Murakami, K., Taguchi, H.: Gesture recognition using recurrent neural networks. In: Proceedings of the Conference on Human Factors in Computing Systems, pp. 237–242 (1991)

4. Dipietro, L., Sabatini, A.M., Dario, P.: A survey of glove-based system their applications. IEEE Trans. Syst. Man Cybern. Part C Appl. Rev. **38**(4), 461–482 (2008)
5. Cabrera, M.E., Bogado, J.M., Fermin, L., Acuña, R., Ralev, D.: Glove-based gesture recognition system. In: Proceedings of the 15th International Conference on Climbing and Walking Robots and the Support Technologies for Mobile Machines, Baltimore, MD, USA, pp. 747–753 (2012)
6. Phi, L.T., Bui, T.T.Q., Nguyen, H.D., Vu, T.T.: A glove-based gesture recognition system for Vietnamese sign language. In: 15th International Conference on Control, Automation and Systems (ICCAS 2015), October 13–16, BEXCO, Busan, Korea (2015)
7. Yang, R., Sarkar, S., Loeding, B.L.: Handling movement epenthesis and hand segmentation ambiguities in continuous sign language recognition using nested dynamic programming. IEEE Trans. Pattern Anal. Mach. Intell. **32**(3), 462–477 (2010)
8. Mehdi, S.A., Khan, Y.N.: Sign language recognition using sensor gloves. In: Proceedings of the 9th International Conference on Neural Information Processing, vol. 5, pp. 2204–2209 (2002)
9. Bui, T.T.Q., Hong, K.-S.: Evaluating a color-based active basis model for object recognition. Comput. Vis. Image Underst. **116**(11), 1111–1120 (2012)
10. Bui, T.D., Nguyen,L.T.: Recognition of Vietnamese sign language using mems accelerometers. In: Proceedings of the First International Conference on Sensing Technology, Palmerston North, New Zealand, November 21–23, pp. 118–122 (2005)
11. Bui, T.D., Nguyen, L.T.: Recognition postures in Vietnamese sign language using mems accelerometers. IEEE Sens. J. **7**(5), 707–712 (2007)
12. Gao, W., Ma, J., Wu, J., Wang, C.: Sign language recognition based on HMM/ANN/DP. Int. J. Pattern Recognit. Artif. Intell. **14**(5), 587–602 (2000)
13. http://www.who.int/mediacentre/factsheets/fs300/en/

Stochastic Bounds for Inference in Topic Models

Xuan Bui[1,2(✉)], Tu Vu[1], and Khoat Than[1]

[1] Hanoi University of Science and Technology,
No.1, Dai Co Viet Road, Hanoi, Vietnam
thanhxuan1581@gmail.com, vutu201130@gmail.com,
khoattq@soict.hust.edu.vn
[2] Thai Nguyen University of Information and Communication Technology,
Quyet Thang, Thai Nguyen, Vietnam

Abstract. Topic models are popular for modeling discrete data (e.g., texts, images, videos, links), and provide an efficient way to discover hidden structures/semantics in massive data. The problem of posterior inference for individual texts is particularly important in streaming environments, but often intractable in the worst case. Some existing methods for posterior inference are approximate but do not have any guarantee on neither quality nor convergence rate. Online Maximum a Posterior Estimation algorithm (OPE) [13] has more attractive properties than existing inference approaches, including theoretical guarantees on quality and fast convergence rate. In this paper, we introduce three new algorithms to improve OPE (so called OPE1, OPE2, OPE3) by using stochastic bounds when doing inference. Our algorithms not only maintain the key advantages of OPE but often outperform OPE and existing algorithms. Our new algorithms have been employed to develop new effective methods for learning topic models from massive/streaming text collections.

Keywords: Topic models · Posterior inference · Online map estimation · Large-scale learning · OPE

1 Introduction

Latent Dirichlet Allocation (LDA) [4] is the class of Bayesian networks. It has found successful applications in a wide range of areas including text modeling [5], bioinformatics [2,9], history [6], politics [12,17], psychology [8]. Estimation of posterior distributions for individual documents is one of the core issues in LDA. Recently, this estimation problem is considered by many researchers, and many methods such as Variational Bayes (VB) [4], Collapsed Variational Bayes (CVB) [19], CVB0 [1], and Collapsed Gibbs Sampling (CGS) [7,18], OPE [13], have been proposed. The quality of LDA in practice is determined by the quality of the inference method being employed. However, except OPE, none of the mentioned methods has a theoretical guarantee on quality or convergence rate.

Our first contribution is the introduction of some new algorithms for doing posterior inference of topic mixture in LDA by improving OPE algorithm.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_62

They are called OPE1, OPE2, OPE3. The posterior inference problem is in fact nonconvex and is NP-hard [3]. OPE is stochastic in nature and theoretically converges to a local maximal/stationary point of the inference problem. In [13], Than and Doan proved that OPE converges at a rate of $O(1/T)$, which surpasses the best rate of existing stochastic algorithms for nonconvex problems [10,16], where T is the number of iterations. One main drawback of OPE is that there is no guarantee for OPE to get rid of saddle points of the inference problems.[1] In this paper, we propose three new variants of OPE, which are called OPE1, OPE2, and OPE3 for doing inference in topic models. Note that OPE does inference by maximizing an objective function of the inference problem. In our algorithms, we use both upper and lower bounds of the objective function. The usage of both bounds is stochastic in nature and help us reduce the possibility of getting stuck at a local stationary point. Therefore, our new variants are more beneficial than the original OPE [13].

Our second contribution in this paper is that we introduce six new methods for learning LDA from massive/streaming text collections. Those methods employ OPE1, OPE2, and OPE3 as their core routines to do inference. From extensive experiments on two large corpora we find that some of our methods can reach state-of-the-art performance in both predictiveness and model quality.

Organization: The rest of this paper is organized as follows. We introduce an overview of posterior inference with LDA, followed by OPE algorithm in Sect. 2. In Sect. 3, three new algorithm for inference are proposed in detail. Section 4 is application of proposed algorithms applied to online learning LDA. In Sect. 5, we give some results test with large datasets and we conclude the paper in Sect. 6.

Notation: Throughout the paper, we use the following conventions and notations. Bold faces denote vectors or matrices. The unit simplex in the n-dimensional Euclidean space is denoted as $\Delta_n = \{\boldsymbol{x} \in \mathbb{R}^n : \boldsymbol{x} \geq 0, \sum_{k=1}^{n} x_k = 1\}$, and its interior is denoted as $\overline{\Delta}_n$. We will work with text collections with V dimensions (dictionary size). Each document $\boldsymbol{d}$ will be represented as a frequency vector, $\boldsymbol{d} = (d_1, ..., d_V)^T$ where d_j represents the frequency of term j in $\mathbf{d}$. Denote n_d as the length of $\boldsymbol{d}$, i.e., $n_d = \sum_j d_j$.

2 Background on Topic Models and Posterior Inference

A topic model often assumes that a corpus is composed from K topics, $\beta = (\beta_1, .., \beta_K)$. Each document $\mathbf{d}$ is a mixture of those topics and is assumed to arises from the following generative process. For the n^{th} word of $\mathbf{d}$:

- draw topic index $z_{dn}|\theta_d \sim Multinomial(\theta_d)$
- draw word $w_{dn}|z_{dn}, \beta \sim Multinomial(\beta_{z_{dn}})$

[1] A saddle point is not always a (local) maximal point. Further, the inference might have exponentially large number of saddle points.

Each topic mixture $\theta_d = (\theta_1, \ldots, \theta_K)$ represents the contributions of topics to document $\mathbf{d}$, while β_{kj} shows the contribution of term j to topic k. Note that $\theta \in \Delta_K, \beta_k \in \Delta_V, \forall k$. Both θ and $\mathbf{z}$ are hidden variables and are local for each document.

The generative process above generally describes what probabilistic latent semantic analysis (PLSA) is [4,5]. Latent Dirichlet allocation (LDA) further assumes that θ and β are samples of some Dirichlet distributions. More specifically, $\theta \sim Dirichlet(\alpha)$ and $\beta_k \sim Dirichlet(\eta)$ for any topic. The problem of posterior inference for each document $\mathbf{d}$, given a model $\{\beta, \alpha\}$, is to estimate the full joint distribution $p(\mathbf{z}_d, \theta, \mathbf{d}|\beta, \alpha)$. Direct estimation of this distribution is intractable, i.e., NP-hard in the worst case. Hence, existing inference approaches use different schemes. VB, CVB and CVB0 try to estimate the distribution by maximizing a lower bound of the likelihood $p(\mathbf{d}|\beta, \alpha)$, whereas CGS tries to estimate $p(z|d, \beta, \alpha)$. We consider the MAP estimation of topic mixture for a given document d

$$\theta^* = \mathbf{argmax}_{\theta \in \overline{\Delta}_K} Pr(\theta, \mathbf{d}|\beta, \alpha) = \mathbf{argmax}_{\theta \in \overline{\Delta}_K} Pr(d|\theta, \beta) Pr(\theta|\alpha) \tag{B.1}$$

Remember that the density of the K-dimensional Dirichlet distribution with parameter α is $P(\theta|\alpha) \propto \prod_{k=1}^{K} \theta_k^{\alpha-1}$. Therefore problem (B.1) is equivalent to the following:

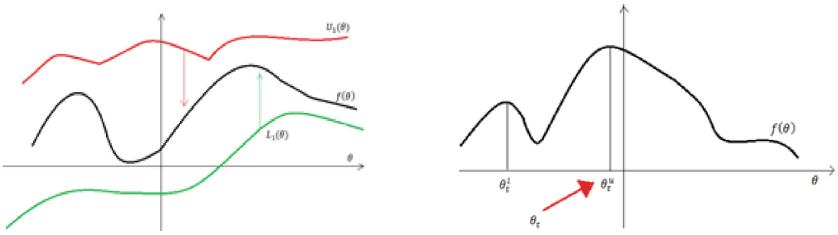
$$\theta^* = \mathbf{argmax}_{\theta \in \overline{\Delta}_K} \sum_j d_j \log \sum_{k=1}^{K} \theta_k \beta_{kj} + (\alpha - 1) \sum_{k=1}^{K} \log \theta_k \tag{B.2}$$

Sontag and Roy [3] showed that this problem is NP-hard in the worst case when $\alpha < 1$. In the case of $\alpha \geq 1$, one can easily show that the problem (B.2) is concave, and therefore it can be solved in polynomial time. Unfortunately, in practice of LDA, the parameter α is often small, says $\alpha < 1$, causing (B.2) to be nonconcave. That is the reason for why (B.2) is intractable in the worst case.

Than and Doan [13] present OPE for doing inference of topic mixtures for documents. The idea of OPE is quite simple. It solves problem (B.2) by iteratively find a vertex of $\overline{\Delta}_K = \{x \in R^K : \sum_{k=1}^{K} x_k \geq \varepsilon > 0\}$ to improve its solution. A good vertex at each iteration is decided by assessing stochastic approximations to the gradient of the objective function. When the number of iterations goes to infinity, OPE will approach to a local maximal/stationary point of problem (B.2).

3 Three New Algorithms for Inference

In this section, we describe important characters of OPE, which were investigated by Than and Doan [13]. Then, we analyze some new perspectives of OPE which lead to our improvements. OPE is a stochastic algorithm. Denote $g_1(\theta) = \sum_j d_j \log \sum_{k=1}^{K} \theta_k \beta_{kj}$ and $g_2(\theta) = (\alpha - 1) \sum_{k=1}^{K} \log \theta_k$. It constructs a sequence of random functions that approximate the objective function of interest



(a) Upper bound and lower bound of objective function

(b) Choose the best point

Fig. 1. Ideas to improve OPE

Algorithm 1. OPE1: The first variant of OPE

Input: document d and model $\{\boldsymbol{\beta}, \alpha\}$
Output: $\boldsymbol{\theta}$ that maximizes
$f(\boldsymbol{\theta}) = \sum_j d_j \log \sum_{k=1}^K \theta_k \beta_{kj} + (\alpha - 1) \sum_{k=1}^K \log \theta_k$.
Initialize $\boldsymbol{\theta_1}$ arbitrary in Δ_K
$f_1^u := \sum_j d_j \log \sum_{k=1}^K \theta_k \beta_{kj}$; $f_1^l := (\alpha - 1) \sum_{k=1}^K \log \theta_k$
for $t = 1, 2, ...\infty$ **do**
Pick f_t^u unifromly from $\{\sum_j d_j \log \sum_{k=1}^K \theta_j \beta_{kj}; (\alpha - 1) \sum_{k=1}^K \log \theta_k\}$
$U_t := \frac{2}{t} \sum_{h=1}^t f_h^u$
$\boldsymbol{e}_t^u := argmax_{\boldsymbol{x} \in \Delta_K} < U_t^{'}(\boldsymbol{\theta}_t), \boldsymbol{x} >$
$\boldsymbol{\theta}_{t+1}^u := \boldsymbol{\theta}_t + \frac{\boldsymbol{e}_t^u - \boldsymbol{\theta}_t}{t}$
Pick f_t^l unifromly from $\{\sum_j d_j \log \sum_{k=1}^K \theta_j \beta_{kj}; (\alpha - 1) \sum_{k=1}^K \log \theta_k\}$
$L_t := \frac{2}{t} \sum_{h=1}^t f_h^l$
$\boldsymbol{e}_t^l := argmax_{\boldsymbol{x} \in \Delta_K} < L_t^{'}(\boldsymbol{\theta}_t), \boldsymbol{x} >$
$\boldsymbol{\theta}_{t+1}^l := \boldsymbol{\theta}_t + \frac{\boldsymbol{e}_t^l - \boldsymbol{\theta}_t}{t}$
$\boldsymbol{\theta}_{t+1} :=$ pick uniformly from $\{\boldsymbol{\theta}_{t+1}^u, \boldsymbol{\theta}_{t+1}^l\}$
end for

$f(\theta) = g_1(\theta) + g_2(\theta)$. OPE works by alternatively choosing $g_1(\theta), g_2(\theta)$ in each iteration to create a random function $F_t(\theta)$. Then, $F_t(\theta)$ converges to $f(\theta)$ when $t \to \infty$.

We recognized that $g_1(\theta) < 0, g_2(\theta) > 0$. Hence, in the first iteration, if we choose $g_1(\theta)$ then $F_1(\theta) < f(\theta)$, the sequence of random functions goes from below $f(\theta)$ according to coordinate. In constrast, if we choose $g_2(\theta)$ in the first iteration then $F_1(\theta) > f(\theta)$, the sequence of random functions goes from above $f(\theta)$. We got an idea to create two sequences of random functions that both converging to $f(\theta)$, one begins with $g_1(\theta)$, other begins with $g_2(\theta)$, both converge to $f(\theta)$ (Fig. 1a). Two random sequences give us more information about objective function, so that we can get a better result when seeking maximal point of $f(\theta)$. OPE1 is aimed at increasing randomness of a stochastic algorithm. Getting idea from random forest, which construct a lot of random trees to get the average result of all trees, we use randonness to create a plenty of choices in

our algorithm. We hope that with fully randomness OPE1 can jump over local stationary points to get higher local stationany point. We pick θ_t uniformly from $\{\theta_t^u, \theta_t^l\}$.

Algorithm 2. OPE2: The second variant of OPE

Input: document d and model $\{\boldsymbol{\beta}, \alpha\}$
Output: $\boldsymbol{\theta}$ that maximizes

$$f(\boldsymbol{\theta}) = \sum_j d_j \log \sum_{k=1}^K \theta_k \beta_{kj} + (\alpha - 1) \sum_{k=1}^K \log \theta_k.$$

Initialize $\boldsymbol{\theta_1}$ arbitrary in Δ_K
$f_1^u := \sum_j d_j \log \sum_{k=1}^K \theta_k \beta_{kj}$; $f_1^l := (\alpha - 1) \sum_{k=1}^K \log \theta_k$
for $t = 1, 2, ...\infty$ **do**
  Pick f_t^u unifromly from $\{\sum_j d_j \log \sum_{k=1}^K \theta_j \beta_{kj}; (\alpha - 1) \sum_{k=1}^K \log \theta_k\}$
  $U_t := \frac{2}{t} \sum_{h=1}^t f_h^u$
  $\boldsymbol{e}_t^u := argmax_{\boldsymbol{x} \in \Delta_K} < U_t^{'}(\boldsymbol{\theta}_t), \boldsymbol{x} >$
  $\boldsymbol{\theta}_{t+1}^u := \boldsymbol{\theta}_t + \frac{\boldsymbol{e}_t^u - \boldsymbol{\theta}_t}{t}$
  Pick f_t^l unifromly from $\{\sum_j d_j \log \sum_{k=1}^K \theta_j \beta_{kj}; (\alpha - 1) \sum_{k=1}^K \log \theta_k\}$
  $L_t := \frac{2}{t} \sum_{h=1}^t f_h^l$
  $\boldsymbol{e}_t^l := argmax_{\boldsymbol{x} \in \Delta_K} < L_t^{'}(\boldsymbol{\theta}_t), \boldsymbol{x} >$
  $\boldsymbol{\theta}_{t+1}^l := \boldsymbol{\theta}_t + \frac{\boldsymbol{e}_t^l - \boldsymbol{\theta}_t}{t}$

  $\boldsymbol{\theta}_{t+1} := \boldsymbol{\theta}_{t+1}^u$ with probability $\frac{exp(f(\boldsymbol{\theta}_{t+1}^u))}{exp(f(\boldsymbol{\theta}_{t+1}^u)) + exp(f(\boldsymbol{\theta}_{t+1}^l))}$
and $\boldsymbol{\theta}_{t+1} := \boldsymbol{\theta}_{t+1}^l$ with probability $\frac{exp(f(\boldsymbol{\theta}_{t+1}^l))}{exp(f(\boldsymbol{\theta}_{t+1}^u)) + exp(f(\boldsymbol{\theta}_{t+1}^l))}$
end for

OPE2 continues with ideas of rising randomness, we pick θ_t from $\{\theta_t^u, \theta_t^l\}$ with probabilities which depend on the value $\{f(\theta_t^u), f(\theta_t^l)\}$. This is smoother than probabilities $\{0.5, 0.5\}$ in OPE1.

The last improvement of OPE, we mix idea with greedy algorithm. We are maximizing a function, so in each iteration, we have two choices for a target point. We get the point that makes the value of objective function is higher (Fig. 1b). The idea of the algorithm OPE3 described our idea. OPE3 works differently from OPE. OPE just constructs one sequence of numbers while OPE3 makes three ones depending on each others. Therefore, the structure of sequence $\{\theta_t\}$ is changed. However, OPE3's properties are similar to OPE's.

Theorem 1.[2] *(**Convergence of OPE algorithms**): Consider the objective function $f(\theta)$ in problem (B.2), given fixed $\boldsymbol{d}, \beta, \alpha$. For OPE1, OPE2, OPE3, the followings hold:*

1. *For any $\theta \in \Delta_K$, $U_t(\theta)$ and $L_t(\theta)$ converges to $f(\theta)$ as $t \longrightarrow +\infty$,*
2. *θ_t converges in probability to a local maximal/stationary point of $f(\theta)$ at a rate of $O(1/t)$.*

[2] The detailed proof of this theorem will be presented in another paper.

Algorithm 3. OPE3 : The third variant of OPE

Input: document d and model $\{\boldsymbol{\beta}, \alpha\}$

Output: $\boldsymbol{\theta}$ that maximizes

$$f(\boldsymbol{\theta}) = \sum_j d_j \log \sum_{k=1}^{K} \theta_k \beta_{kj} + (\alpha - 1) \sum_{k=1}^{K} \log \theta_k.$$

Initialize $\boldsymbol{\theta_1}$ arbitrary in Δ_K

$f_1^u := \sum_j d_j \log \sum_{k=1}^{K} \theta_k \beta_{kj}$; $f_1^l := (\alpha - 1) \sum_{k=1}^{K} \log \theta_k$

for $t = 1, 2, ...\infty$ **do**

Pick f_t^u unifromly from $\{\sum_j d_j \log \sum_{k=1}^{K} \theta_j \beta_{kj}; (\alpha - 1) \sum_{k=1}^{K} \log \theta_k\}$

$U_t := \frac{2}{t} \sum_{h=1}^{t} f_h^u$

$\boldsymbol{e}_t^u := argmax_{\boldsymbol{x} \in \Delta_K} < U_t^{'}(\boldsymbol{\theta}_t), \boldsymbol{x} >$

$\boldsymbol{\theta}_{t+1}^u := \boldsymbol{\theta}_t + \frac{\boldsymbol{e}_t^u - \boldsymbol{\theta}_t}{t}$

Pick f_t^l unifromly from $\{\sum_j d_j \log \sum_{k=1}^{K} \theta_j \beta_k j; (\alpha - 1) \sum_{k=1}^{K} \log \theta_k\}$

$L_t := \frac{2}{t} \sum_{h=1}^{t} f_h^l$

$\boldsymbol{e}_t^l := argmax_{\boldsymbol{x} \in \Delta_K} < L_t^{'}(\boldsymbol{\theta}_t), \boldsymbol{x} >$

$\boldsymbol{\theta}_{t+1}^l := \boldsymbol{\theta}_t + \frac{\boldsymbol{e}_t^l - \boldsymbol{\theta}_t}{t}$

$\boldsymbol{\theta}_{t+1} := argmax_{\boldsymbol{\theta} \in \{\boldsymbol{\theta}_{t+1}^u, \boldsymbol{\theta}_{t+1}^l\}} f(\boldsymbol{\theta})$

end for

4 New Approaches for Learning LDA

ML-OPE and Online-OPE are two examples for exploiting OPE algorithm proposed by Than and Doan [13]. These algorithms learn parameters of the LDA model in online schema and streaming data. The core of two algorithms is OPE, which makes ML-OPE and Online-OPE be state-of-the-art algorithms in learning LDA. They are the best algorithm in comparison to others. Their properties were carefully explored in [13].

Based on evolving the inference method of ML-OPE and Online-OPE by OPE1, OPE2 and OPE3, we obtain six new learning algorithms for LDA so called ML-OPE1, ML-OPE2, ML-OPE3, Online-OPE1, Online-OPE2 and Online-OPE3. Thank to rapid convergence of OPE, the result of algorithm is not affected by the number of iterations T [13] (for example, log predictive probability measurements at $T = 20$ and $T = 100$ are asympotic). Therefore, although the number of iteration of methods OPE1, OPE2 and OPE3 is doublely larger than OPE, this does not affect the runtime of inference algorithm. From Theorem 1, we find out that the convergence rate of OPE3 is equal to the one of OPE [13].

ML-OPE1, ML-OPE2, Online-OPE1 and Online-OPE2 algorithms obtain lots of random properties. Inspired by the supervised learning method-Random forest method (By creating many decision trees which have high random nature and taking the average results of these trees, the method is effective in classification). We then make the idea for creating more random natures of these algorithms, and the inference problem (B.2) can jump out of a local maximum to reach close the global one.

ML-OPE3 and Online-OPE3 use the idea of greedy methods, and get better value at each iteration. The greedy algorithms get the best value at every iteration, but the final outcome can not be optimal. OPE3 tries to find maximum by iteration and selects the best optimal point from two ones, so that the value of objective function is higher in each iteration. This is very effective in the case of large-scale machine learning problems. The OPE3 algorithm derived from OPE2 can consider the probability of choosing best point by 1. OPE3 not only archives stable and fast convergence as OPE but also finds better maximum points. So, OPE3 will change the quality of ML-OPE and Online-OPE.

5 Empirical Results

In this section, we investigate the practical performance of our new variants. Since OPE, OPE1, OPE2, OPE3 can play a role in the core subroutine of large-scale learning methods for LDA, we will evaluate the inference algorithms through ML-OPE and Online-OPE based on changing their core inference. By this way, we also see how helpful our new algorithms for posterior inference are. So, thank to changing OPE by our three algorithms in ML-OPE and Online-OPE, we get six new methods for learning LDA. Our simulation results show comparison between the proposed algorithms and the previous ones.

The following two large corpora were used in our experiments: Pubmed consists of 330,000 articles from the Pubmed central; New York Times consists of 300,000 news.[3] To avoid randomness, on each dataset one of learning methods is run for 5 times and reported its average results.

Parameter Settings:

- *Model parameters:* $K = 100, \alpha = \frac{1}{K}, \eta = \frac{1}{K}$
- *Inference parameters:* $T = 20$
- *Learning parameters:* minibatch size $S = |C_t| = 5000, \kappa = 0.9, \tau = 1$

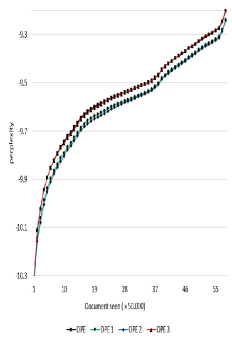
Than and Doan [13] observed that these parameter adapted best for ML-OPE and Online-OPE.

Performance Mesures: We use *NPMI* [11] and *Predictive Probability* [14] to evaluate these learning methods. Log predictive probability measures the predictiveness and generalization of a model to new data, while NPMI evaluates semantics quality of an individual topic in these models.

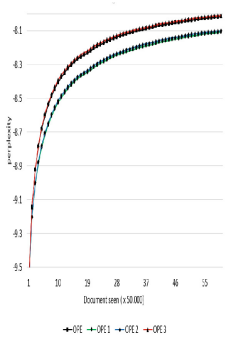
From Figs. 2 and 3, we can see that improved OPE1 and OPE2 making ML-OPE and Online-OPE work worse in both 2 measures, while OPE3 works well with Online-OPE.

Variant of OPE aims to seek θ that maximizes a function $f(\theta)$ on a simplex using stochastic bounds. ML-OPE then updates model parameter β and Online-OPE gets variational element λ. The quality of θ found by OPE directly affects to the quality of $\{\beta, \lambda\}$, then the measures. In practice, OPE performs an fast

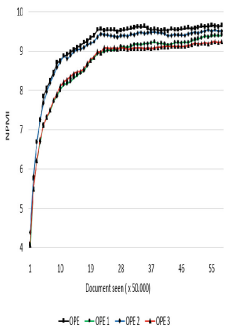
[3] The datasets were taken from http://archive.ics.uci.edu/ml/datasets.



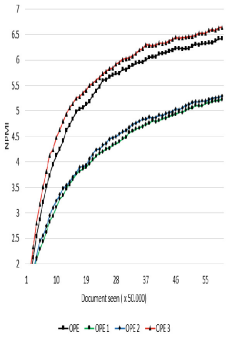
(a) On NYT



(b) On Pubmed



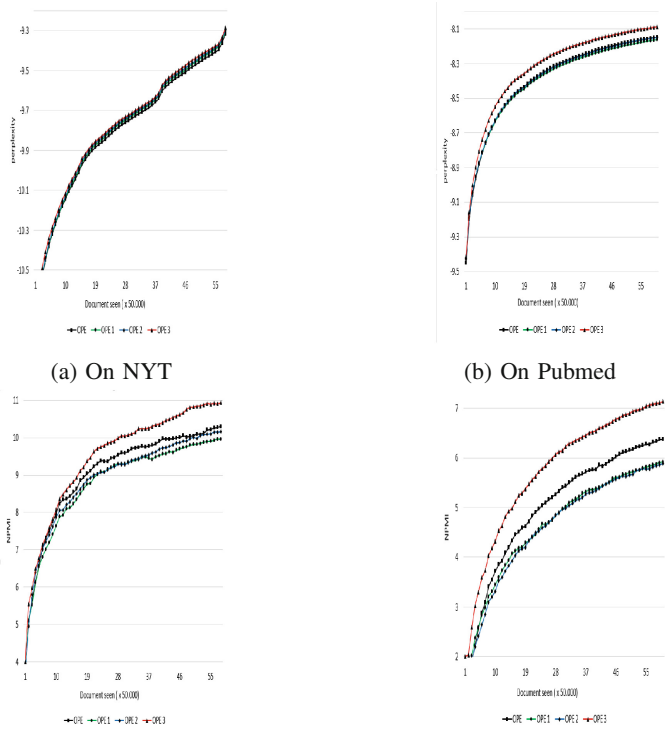
(c) On NYT



(d) On Pubmed

Fig. 2. Performance of ML-OPE when its core inference routine is done by different methods. Note that OPE3 often performs comparably or better than other inference methods.

convergence and stable character. Stability is shown by the number of iterations T. Just $T = 20$ iterations for OPE results in the same predictiveness level as $T = 100$. So, OPE converges very fast. OPE is also stable. Than and Doan did experiments by running 10 times OPE and observed that the results were not too different. OPE1 and OPE2 increase randomness of OPE, but the results of ML-OPE and Online-OPE are not better. With converge property of OPE, it suggests that randomly choose θ_t from $\{\theta_t^u, \theta_t^l\}$ make worse θ from some first iterations, therefore OPE converges fast to worse results. The last improvement makes ML-OPE and Online-OPE work more efficient. This demonstrates our idea of using two random sequences of functions to approximate a objective function. The idea of increasing the randomness and greedy of algorithm are exploited here. Firstly, two random sequences of function are used to rise our participants and information relevant to objective function. Hence in the next iteration we have more choices in θ_t. Secondly, choosing θ_t from $\{\theta_t^u, \theta_t^l\}$ that makes the value of $f(\theta)$ higher in each iteration came from idea of greedy algorithms. There are many ways of choices here, but we designed a best way to create θ_t from $\{\theta_t^u, \theta_t^l\}$.



(a) On NYT (b) On Pubmed

(c) On NYT (d) On Pubmed

Fig. 3. Performance of Online-OPE when its core inference routine is done by different methods. It is easy to see that OPE3 often outperforms the others.

This approach is simple, because it does not need extra parameter. Of course, we can compose θ_t from $\{\theta_t^u, \theta_t^l\}$ by linear combination or something like that. But these approaches make model more complicated. In our experiment, we chose $\theta_t = \alpha\theta_t^u + (1-\alpha)\theta_t^l$ and we get the results better. But we increased the number of parameters in model and we had to choose α empirically. Choosing θ_t in improved OPE3 is more efficient and simpler. OPE can be exploited with general function $f(x)$ if $f(x)$ is a form of $g_1(x) + g_2(x)$. In this paper, we make OPE work better by choose the kind of $g_1(x)$ and $g_2(x) : g_1(x) \leq 0$, $g_2(x) \geq 0$. So for other work using improved OPE3, we can choose the same function $g_1(x), g_2(x)$ and get the better result.

6 Conclusion

We have discussed how posterior inference for individual texts in topic models can be done efficiently. Our novel algorithms (OPE1, OPE2, OPE3) have a theoretical guarantee on quality and convergence rate. In practice, OPE3 can

do inference very fast and effective, and can be easily extended to a wide class of probabilistic models. By exploiting OPE1, OPE2, OPE3 carefully, we have arrived at six efficient methods for learning LDA from data streams or large corpora. As a result, they are good candidates to help us deal with text streams and big data.

Acknowledgments. This research is funded by Vietnam National Foundation for Science and Technology Development (NAFOSTED) under Grant Number 102.05-2014.28 and by the Air Force Office of Scientific Research (AFOSR), Asian Office of Aerospace Research & Development (AOARD), and US Army International Technology Center, Pacific (ITC-PAC) under Award Number FA2386-15-1-4011.

References

1. Asuncion, A., Welling, M., Smyth, P., Teh, Y.W.: On smoothing and inference for topic models. In: Proceedings of the Twenty-Fifth Conference on Uncertainty in Artificial Intelligence, pp. 27–34 (2009)
2. Liu, B., Liu, L., Tsykin, A., Goodall, G.J., Green, J.E., Zhu, M., Kim, C.H., Li, J.: Identifying functional miRNA-mRNA regulatory modules with correspondence latent Dirichlet allocation. Bioinformatics **26**(24), 3105 (2010)
3. Sontag, D., Roy, D.M.: Complexity of inference in latent Dirichlet allocation. In: Advances in Neural Information Processing Systems (NIPS) (2011)
4. Blei, D.M., Ng, A.Y., Jordan, M.I.: Latent Dirichlet allocation. J. Mach. Learn. Res. **3**(3), 993–1022 (2003)
5. Blei, D.M.: Probabilistic topic models. Commun. ACM **55**(4), 77–84 (2012)
6. Mimno, D.: Computational historiography: data mining in a century of classics journals. J. Comput. Cult. Heritage **5**(1), 3 (2012)
7. Mimno, D., Hoffman, M.D., Blei, D.M.: Sparse stochastic inference for latent Dirichlet allocation. In: Proceedings of the 29th Annual International Conference on Machine Learning (2012)
8. Schwartz, H.A., Eichstaedt, J.C., Dziurzynski, L., Kern, M.L., Blanco, E., Kosinski, M., Stillwell, D., Seligman, M.E.P., Ungar, L.H.: Toward personality insights from language exploration in social media. In: AAAI Spring Symposium Series (2013)
9. Pritchard, J.K., Stephens, M., Donnelly, P.: Inference of population structure using multilocus genotype data. Genetics **155**(2), 945–959 (2000)
10. Mairal, J.: Stochastic majorization-minimization algorithms for large-scale optimization. In: Advances in Neural Information Processing Systems, pp. 2283–2291 (2013)
11. Lau, J.H., Newman, D., Baldwin, T.: Machine reading tea leaves: automatically evaluating topic coherence and topic model quality. In: Proceedings of the Association for Computational Linguistics, pp. 530–539 (2014)
12. Grimmer, J.: A Bayesian hierarchical topic model for political texts: measuring expressed agendas in senate press releases. Polit. Anal. **18**(1), p1–35 (2010)
13. Than, K., Doan, T.: Guaranteed algorithms for inference in topic models (2015). arXiv preprint arXiv: 1512.03308
14. Hoffman, M.D., Blei, D.M., Wang, C., Paisley, J.: Stochastic variational inference. J. Mach. Learn. Res. **14**(1), 1303–1347 (2013)
15. Aletras, N., Stevenson, M.: Evaluating topic coherence using distributional semantics. In: Proceedings of the 10th International Conference on Computational Semantics, pp. 13–22 (2013)

16. Ghadimi, S., Lan, G.: Stochastic first-and zeroth-order methods for nonconvex stochastic programming. SIAM J. Optim. **23**(4), 2341–2368 (2013)
17. Gerrish, S., Blei, D.: How they vote: issue-adjusted models of legislative behavior. Adv. Neural Inf. Process. Syst. **25**, 2762–2770 (2012)
18. Griffiths, T.L., Steyvers, M.: Finding scientific topics. In: Proceedings of the National Academy of Sciences of the United States of America (2004)
19. Teh, Y.W., Newman, D.M., Welling, A.: Collapsed variational Bayesian inference algorithm for latent Dirichlet allocation. In: Advances in Neural Information Processing Systems, vol. 19, pp. 13–53 (2007)

Swarm Intelligence-Based Approach for Macroscopic Scale Odor Source Localization Using Multi-robot System

Anh-Quy Hoang(✉) and Minh-Trien Pham(✉)

VNU University of Engineering and Technology,
144 Xuan Thuy, Cau Giay, Hanoi, Vietnam
{quyha,trienpm}@vnu.edu.vn

Abstract. Odor source localization is a problem of great importance. Two mainstream methods among numerous proposed ones are probabilistic algorithms and bio-inspired algorithms. Compared to probabilistic algorithms, biomimetic approaches are much less intensive in term of computational cost. Thus, despite their slightly worse performance, biomimetic approaches have received much more attention. In this paper, a novel method based on a bio-inspired algorithm - Particle Swarm Optimization (PSO) - is proposed for a multi-robot system (MRS). The proposed algorithm makes use of wind information and immediate odor gradient to enhance the performance of the MRS. A mechanism based on Artificial Potential Field (APF) is utilized to ensure non-collision movement of the robots. This method is tested by simulation on Matlab. Data for the test scenarios, all in large scales, are generated using Fluent. Nearly 2000 runs are carried out and the simulation results confirm the proposed algorithm's effectiveness.

Keywords: PSO · Odor source localization · MRS · APF

1 Introduction

Odor source localization is indispensable in various essential tasks. Among those tasks are tracking down drugs, searching for bombs or explosives, mitigating industrial disasters by finding leakages of dangerous substances, locating victims under debris, etc. Some of the tasks could not, or should not, be performed by trained animal because they could not survive the possible danger, or the tasks are too physically demanding for an animal. Thus, we should develop artificial systems, specifically robotic systems, to replace animals in such tasks. In comparison with trained animals, robots equipped with olfactory sensing systems have a number of advantages. They could be used in large numbers and work in extreme conditions, with no training or care. Furthermore, if a robot is destroyed in a mission (e.g. localizing mines), the problem is not as big as in the case of a trained dog.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_63

1.1 Robotic Odor Localization

The problem of odor source localization is, in essence, an optimization problem. As the source releases odor, there is a certain distribution of odor concentration in the search space. The goal then becomes locating the area of highest odor concentration. Nevertheless, efforts to apply classical search or optimization methods have been hindered by the special characteristics of odor plumes. The odor plumes are easily influenced by turbulence, so they could be meandering and intermittent, i.e. the concentration map does not remain the same during the search. The turbulent structure of air flow over a large space may also cause irregular structure with conterminous high and low concentrations. Furthermore, prolonged local maxima could possibly exist under certain conditions. As classic methods do not work, there is a need for modifying them or devising new methods to tackle the task of odor source localization.

The odor source localization task consists of three stages. The first stage is plume detection. In this stage, the robots look for the presence of a chemical (usually in the form of a plume) with no information regarding the position of the source or any other hints. The primary method used is random walk, in which the robots move randomly to detect clues of the plumes. Besides, some authors also deploy different strategies to enhance the global searching performance in the first stage. For instance, Marques et al. [1] used a robot system where robots tend to avoid each other by default. This behavior increases the probability of detecting a plume when there are no chemical cues in the neighborhood. In [2], Jatmiko et al. deployed an algorithm named Charged PSO. Robots operating under this algorithm exert repulsive forces to other in order to avoid collisions and in the meantime improve the system's ability to explore the environment. When a plume is detected, the system turns to the second stage. The second stage is plume tracking, also known as plume traversal, in which the robots have already been in a plume and they try to traverse the plume to approach its source. A significant body of research has focused on the second stage, this is to be discussed further in the next subsection. The final stage is source declaration (or source identification), when the source is declared to be found in the immediate vicinity of the searching agents (robots) [3]. In this stage, most authors declare sources when the signal returned by the olfactory sensing system exceeds a threshold. It is noteworthy that the three stages are not implemented successively because the robots may be out of the plume at any time due to the dynamic nature of odor plumes. In a realistic process, robots continually switch between plume finding and plume traversal.

1.2 Related Studies

The need for research on robotic odor localization is strong, and it is emphasized by the fact that current technology is still behind the capabilities of the animal although first studies are dated back to as early as the 1990s. One of the first studies in this field is [4], where the use of chemically sensitive robots are discussed. Since then, various aspects of the problems have been discussed and

nowadays, this is a mature and popular research area. Reference [5] provides a comprehensive taxonomy and survey in this area.

A number of algorithms for a single robot to traverse a plume to its sources have been proposed with some noble ones: casting algorithm, surge-spiral algorithm and surge-cast algorithm [6]. Although the approaches are effective and applicable to some degree, single robots are not usually used because they are not time-efficient. Instead, multi-robot systems (MRS) are usually the choice in practical applications. The advantages of multi-robot systems over single robots are not limited to time efficiency. Due to their abilities to share information, cooperate and coordinate, multi-robot systems have a higher robustness to local optima, wider coverage and higher degree of accuracy, compared to single robots. It is the reason why recent studies on robotic odor localization are devoted to improving the performance of multi-robot systems. The most widely used methods for multi-robot systems include biomimetic algorithms and probabilistic inference approaches. In [1,7–9], the authors presented novel algorithms based on PSO to track odor plumes. PSO is very effective in static environments, however, in most problems of odor source localization, the environment is dynamic. Therefore, modifications must be made to PSO. In [10–12], odor samples are obtained and used to generate a probability distribution of the source location. The next movement is determined based on this generated probability.

1.3 Main Contributions

In spite of the large body of research, there is still much to be done. There is a variety of environmental conditions, but each study is normally only focused on one of them. Normally, gradient-based algorithms are not used in macroscopic scales. In a large search space, there are areas without any cue of the plume, and the robots within such areas would be unable to be guided by the gradient.

This paper focuses on the task of localizing a single odor source in a large area with low-turbulence using a multi-robot system. Main contributions of this study are: (1) Proposing a modified algorithm based on PSO, which is specialized for the task of odor localization. In the algorithm, APF and gradient-based search are integrated into the conventional PSO, wind information is also utilized to enhance the search performance. (2) Evaluating the effectiveness of the proposed algorithm through numerous simulations in dynamic environments.

2 Theoretical Background and Methodology

2.1 Artificial Potential Field

Proposed by Khatib in 1986 [13] for single robots, nowadays Artificial Potential Field (APF) is widely used in works related to path planning of both single robots and MRS. When APF is applied, each robot will be in a potential field that exerts forces on them to prevent collisions and drive they toward their destinations. Magnitudes of the potential forces are continuously updated, based

on information on the relative position of each robot with other robots and obstacles. APF forces could be attractive or repulsive forces. In this research, APF is used to keep the robots from collisions, so only repulsive forces are generated and they exert on particles within the repulsive region of others. The repulsive force between a particle and another particle or an obstacle is given by:

$$\mathbf{F}_{APFij} = (F_{\max} - k \times r_{ij}^2) \times \frac{\mathbf{r}_{ij}}{r_{ij}} \times (H(0) - H(r_1)) \tag{1}$$

Where $\mathbf{r}_{ij}$ is the distance vector from object j to object i (robot-robot or robot-obstacle), r_{ij} is the module of $\mathbf{r}_{ij}$. F_{max} is the maximum value of the repulsive force. H(x) is Heaviside step function. r_1 is called separation radius, it is the radius of repulsive region surrounding each robot. The separation radius must be smaller than sensing range of a robot. k is not a constant, it is calculated in such a way that $\mathbf{F}_{APFij}$ is zero when i = j or r = r_1. The total repulsive force exerted on i^{th} robot of the system is:

$$\mathbf{F}_{APFi} = \sum_{j=1}^{N} \mathbf{F}_{APFij} \tag{2}$$

Where N is the number of robots. The impact of $\mathbf{F}_{APFi}$ on overall velocity is controlled by F_{max} and k. As F_{max} increases, the particle is less likely to approach obstacles. $\mathbf{v}_{separation}$ is the component velocity that helps the robots to avoid mutual collisions and collisions with obstacles. This component velocity is determined by total repulsive forces on the swarm robots.

2.2 Gradient-Based Search

Gradient-based search is an efficient method, which is most suitable for differentiable, uni-modal functions. Both conditions are not satisfied in the case of odor source location. However, if we divide the search space into many small regions, in each region the fitness function (odor concentration) may be uni-modal and have a smooth gradient, then a gradient search is applicable locally. In regions where the odor concentration is too low and undetectable, the velocity component resulted from gradient-based search is set to zero.

Generally, the gradient search is implemented as described in the following formula:

$$\mathbf{p}_{t+1} = \mathbf{p}_t + \alpha.\nabla f(\mathbf{p}_t) \tag{3}$$

Where $\mathbf{p}_t$ represents the space coordinates at time t, α is a positive constant. In the cases where the function f is not analytically known (as in this application, where the search space is unknown), the gradient can be empirically estimated.

$\mathbf{v}_{grad}$ is defined as the component velocity that guides each robot towards the direction of gradient vector at its position. In the proposed algorithm, to obtain the immediate gradient vector for a robot, first, the odor concentration at positions of all robots are gathered. That data are then used to generated

an interpolated function. The immediate gradient vector at the position of each robot is obtained using central difference method.

2.3 Utilizing Wind Information

Although the odor plumes may be intermittent and meandering, at any given time a robot within the plume could get closer to the source by moving upwind. This does not hold true for all robots because outside of the plumes there may be circulations of air, which lead to nowhere; besides, a robot may also happen to stay upwind of the source. We make use of this property and add an upwind velocity, namely $\mathbf{v}_{uw}$, to the total velocity of a robot. If a robot is in the plume traversal phase, this component velocity is directly opposite to the wind at its position, with the magnitude being proportional to wind velocity; otherwise, this velocity is zero.

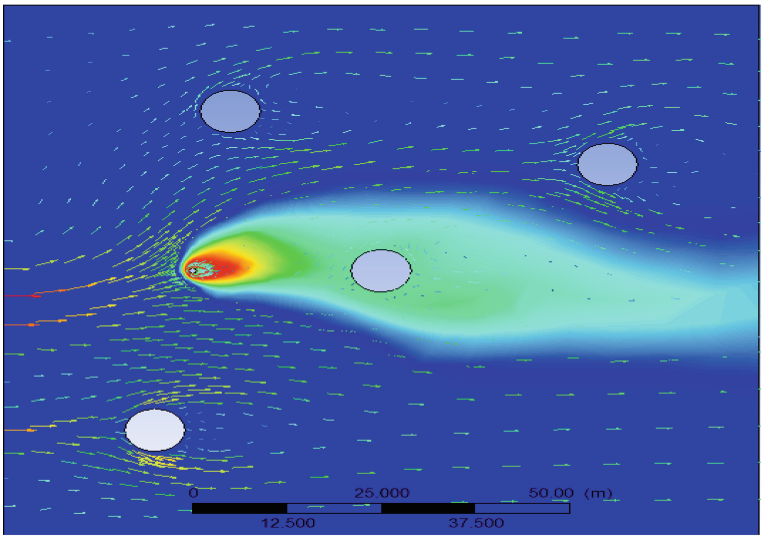


Fig. 1. Scenario 1 with odor concentration map and wind velocity

2.4 The Proposed Algorithm

This subsection contains a brief introduction to PSO and a detailed explanation of the proposed algorithm.

PSO is a stochastic heuristic and optimization algorithm that enables a swarm of homogeneous particles to collectively explore a space for the best solution. Each solution is represented by a point (position) in a multi-dimensional solution space. Particles continuously move toward better solutions by an intelligent mechanism inspired by natural swarm behavior, in which their velocity is

updated regarding the best solutions discovered by each particle (personal best) and by the whole swarm (global best). The velocity of each particle consists of three components that are responsible for guiding the particle toward its personal best position (cognitive velocity), global best position (social velocity) and preventing sudden changes of direction (inertial velocity). Velocities and positions of the particles are updated until stopping criteria are satisfied, according to the following formulas:

$$\mathbf{v}_{inertial} = w \times \mathbf{v}_{t-1} \tag{4}$$

$$\mathbf{v}_{cognitive} = a_1 \times u_1 \times \varphi(\mathbf{p}_{t-1} - \mathbf{x}_{t-1}) \tag{5}$$

$$\mathbf{v}_{social} = a_2 \times u_2 \times \varphi(\mathbf{g}_{t-1} - \mathbf{x}_{t-1}) \tag{6}$$

$$\mathbf{v}_t = \mathbf{v}_{inertial} + \mathbf{v}_{cognitive} + \mathbf{v}_{social} \tag{7}$$

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{v}_t \tag{8}$$

where: $\mathbf{v}_t$: velocity of the swarm, w: inertial factor, a_1: cognitive coefficient, a_2: social coefficient, u_1 and u_2: random number in [0, 1], $\mathbf{p}_t$: personal best positions, $\mathbf{g}_t$: global best positions, $\mathbf{x}_t$: position of the swarm. The subscript t denotes time.

$\varphi(x)$ is a matrix function used to get the distance between positions. In (4) and (5), $\varphi(\mathbf{p}_{t-1} - \mathbf{x}_{t-1})$ and $\varphi(\mathbf{g}_{t-1} - \mathbf{x}_{t-1})$ return the distance vectors from particles to their best positions and to global best position. As a population based algorithm, PSO is suitable to apply to an MRS where each robot is modeled as a particle. The robots' movements resemble those of ideal particles described above, except for that real robots may collide. Thus, additional techniques must be utilized in the actual implementation of PSO for MRS. Moreover, PSO is not effective when the environment is dynamic as in the problem of localizing odor sources. The proposed algorithm is developed to overcome these shortcomings. With the introduction of $\mathbf{v}_{separation}$, $\mathbf{v}_{grad}$, and $\mathbf{v}_{uw}$, the robots are capable of avoiding collisions while moving toward the odor source in an ever changing environment. The stages in the proposed algorithm are basically the same as those of PSO, namely initialization, updating velocities and positions, and checking stopping criteria. The swarm will be in the update phase until they meet stopping criteria. However, in the proposed algorithm, the velocities and positions are updated with new formulas:

$$\mathbf{w} = sig(\mathbf{d} \times k + l) \tag{9}$$

$$\mathbf{v}_{inertial} = \mathbf{w}. \times \mathbf{v}_{t-1} \tag{10}$$

$$\mathbf{v}_{cognitive} = C \times sig(\varphi(\mathbf{p}_{t-1} - \mathbf{x}_{t-1}) \times u + v) \tag{11}$$

$$\mathbf{v}_{social} = S \times sig(\varphi(\mathbf{g}_{t-1} - \mathbf{x}_{t-1}) \times u + v) \tag{12}$$

$$\mathbf{v}_t = \mathbf{v}_{inertial} + \mathbf{v}_{cognitive} + \mathbf{v}_{social} + \mathbf{v}_{separation} + \mathbf{v}_{grad} + \mathbf{v}_{uw} + \mathbf{v}_{rand} \tag{13}$$

$$\mathbf{x}_t = \mathbf{x}_{t-1} + \mathbf{v}_t \tag{14}$$

where **d**: immediate population density at the position of a robot, $\mathbf{v}_{rand}$: random component velocity, .×: element-wise matrix multiplication, k, l, u, v: adjusting parameters, C: maximum cognitive velocity, S: maximum social velocity. $sig(\mathbf{x})$: element-wise sigmoid function on matrix:

$$sig(x)(i,j) = \frac{1}{1 + e^{-x(i,j)}} \tag{15}$$

The main difference between conventional PSO and the proposed algorithm is how velocity is updated. During the exploration, each robot finds itself in an ever-changing potential field. In the proposed algorithm, new velocity components, namely $\mathbf{v}_{separation}$, $\mathbf{v}_{grad}$, and $\mathbf{v}_{uw}$, are introduced. Its inertial factor depends on immediate population density, $\mathbf{v}_{cognitive}$ and $\mathbf{v}_{social}$ are functions of distance to best positions, described by the sigmoid function. The velocities are high at large distance and low at small distance. This reduces the possibility of collision, meanwhile, yields a high performance and enable the swarm to search for odor sources in a dynamic environment.

3 Simulation and Results

3.1 Simulation Setup

In this research, the proposed algorithm is implemented on a homogeneous MRS in Matlab environment. Each robot's size is negligible in comparison with the size of search space. The system has direct communication with unlimited range (the communication range is beyond search space's boundaries). The separation radius r_1 is 5 m. Maximum velocity is 1.5 m/step (the distance between two consecutive positions of a robot could not exceed 1.5 m). Robots are equipped with electronic nose and anemometer to detect the odor molar concentration and get information about the wind at their positions in search space.

The search space size is 100 m × 100 m. In the Cartesian coordinate system with coordinates specified in meters, the ranges of x and y coordinates are both [−50, 50]. The proposed algorithm's effectiveness is evaluated in three scenarios, which are created using $Fluent^{TM}$. In the simulation, all obstacles in the search space are static cylindrical obstacles. The radii of cylindrical obstacles (if any) used in all scenarios are 4 m.

In the first scenario, the odor source is located at (40, −25) and there is no obstacle. The search space is an open square, where the wind blows from the south to the north, the right side and the left side are both walls. The wind direction is normal to search space boundary. Odor emission rate and the speed of air flow in this scenario are constant. This scenario is used to test the ability of the swarm to search for sources that are near the wall.

The search space in the second scenario includes a source at (−25, 0) and four static obstacles at (−30, −30), (−20, 30), (0, 0) and (30, 20) (Fig. 1). Air flows from the left to the right side, with the function describing its velocity more complicated than in the previous scenario. Both the velocity of air flow and

emission rate of odor change over time. The concentration of odor at any given points and the direction of the plume change over time. However, the change is small. This scenario is used to confirm the effectiveness of the algorithm in avoiding collision and finding the target.

In the last scenario, an odor source is placed at (0, −50) and there is no obstacle. The air flows from the left side to the right side. The wind speed and rate of emission are designed so that the plume is intermittent. This scenario is used to assess the swarm's ability to locate odor source when the plume is highly unstable.

3.2 Results and Discussion

In each scenario, population size varies between 5, 8, 11, 14, 17 and 20. The simulation results acquired after 100 runs (for each scenario and each population size) are presented in Tables 1 and 2. The maximum allowed number of iterations is 200, i.e. within 200 iterations, if the swarm does not converge or the threshold value of odor concentration is not reached, the search is deemed to be unsuccessful.

From tables, we could see that both success rate and time consumption are improved by the increase in swarm population. This is understandable, since a larger number of robots stand a higher chance to detect a plume and get more chemical cues. With the method used in this studies, a larger number of robots also give a more accurate estimated gradient. Especially, in scenario 2 and 3, when the number of robots is 8 or higher, all searches returned positive results. It is possible that when the population is large enough, success is guaranteed.

The results presented in the tables also indicate that the algorithm deal with obstacles better than with the intermittence of the plume and with the sources being close to the walls.

In robotic odor localization and the searching task using MRS at large, the time constraint is always crucial. Figure 2 shows how success rate in scenario 1

Table 1. Success rate in different scenarios (%)

Scenario	5 robots	8 robots	11 robots	14 robots	17 robots	20 robots
1	33	56	89	90	93	97
2	93	100	100	100	100	100
3	84	100	100	100	100	100

Table 2. Average time consumption of successful runs (iterations)

Scenario	5 robots	8 robots	11 robots	14 robots	17 robots	20 robots
1	116.67	102.36	84.72	68.39	62.46	61.45
2	89.74	50.06	40.04	35.23	28.47	27.11
3	115.82	64.68	50.17	41.66	35.29	32.79

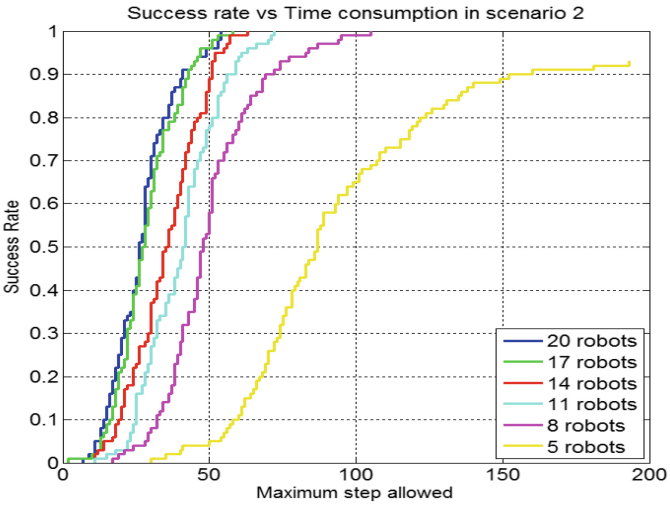


Fig. 2. Success rate vs. Time consumption in scenario 2

may change if the searching time is limited. This is useful for evaluating the algorithm's time efficiency. We could see that, for example, if the searching time is limited to 50 iterations, how the success rate is contingent upon the swarm population. While the success rate stays at as low as 5 % with 5 robots, it increases dramatically to around 65 % and 90 % with 8 and 14 robots, respectively.

4 Conclusion and Future Works

This study is concerned with a new algorithm based on PSO, which could be applied to multi-robot systems to solving the problem of localizing odor sources in large-scale dynamic environments. In this algorithm, the method for generating velocity components of PSO is changed and furthermore, additional velocity components are introduced. Information regarding wind and gradient is exploited for better performance.

Simulation results on Matlab show that the new algorithm is reliable. It can tackle obstacles in the search space very well and could ensure collision-free motions. The robot swarm also perform well even when the plume is intermittent.

In this works, we assume excellent communication, anemometry, odometry and positioning of the robot system, which is not the case in the real word. We will conduct further work on dealing with real-world condition and applying the algorithm on a real MRS.

Acknowledgments. This work has been supported by Vietnam National University, Hanoi (VNU), under Project No. QG.15.25.

References

1. Marques, L., Nunes, U., de Almeida, A.T.: Particle swarm-based olfactory guided search. Auton. Robots **20**(3), 277–287 (2006)
2. Jatmiko, W., Sekiyama, K., Fukuda, T.: A PSO-based mobile robot for odor source localization in dynamic advection-diffusion with obstacles environment: theory, simulation and measurement. IEEE Comput. Intell. Mag. **2**(2), 37–51 (2007)
3. Nakamoto, T. (ed.): Essentials of Machine Olfaction and Taste. Wiley, Chichester (2016)
4. Larcombe, M.H.E.: Robotics in nuclear engineering: computer-assisted teleoperation in hazardous environments with particular reference to radiation fields (1984)
5. Kowadlo, G.: Andrew Russell, R.: Robot odor localization: a taxonomy and survey. Int. J. Robot. Res. **27**(8), 869–894 (2008)
6. Lochmatter, T., Martinoli, A.: Simulation experiments with bio-inspired algorithms for odor source localization in laminar wind flow. In: Seventh International Conference on Machine Learning and Applications, ICMLA 2008. IEEE (2008)
7. Marques, L., de Almeida, A.T.: Finding odours across large search spaces: a particle swarm-based approach. In: Armada, M.A., de González Santos, P. (eds.) Climbing and Walking Robots, pp. 419–426. Springer, Heidelberg (2005)
8. Gong, D.-W., et al.: Modified particle swarm optimization for odor source localization of multi-robot. In: 2011 IEEE Congress of Evolutionary Computation (CEC). IEEE (2011)
9. Gong, D.-W., Zhang, Y., Qi, C.-L.: Localising odour source using multi-robot and anemotaxis-based particle swarm optimisation. IET Control Theory Appl. **6**(11), 1661–1670 (2012)
10. Vergassola, M., Villermaux, E., Shraiman, B.I.: Infotaxis as a strategy for searching without gradients. Nature **445**(7126), 406–409 (2007)
11. Li, J.-G., et al.: Odor source localization using a mobile robot in outdoor airflow environments with a particle filter algorithm. Auton. Robots **30**(3), 281–292 (2011)
12. Li, J.-G., et al.: Odor-source searching using a mobile robot in time-variant airflow environments with obstacles. In: 2014 33rd Chinese Control Conference (CCC). IEEE (2014)
13. Khatib, O.: Real-time obstacle avoidance for manipulators and mobile robots. Int. J. Robot. Res. **5**(1), 90–98 (1986)

The Analyzes of Network-on-Chip Architectures Based on NOXIM Simulator

Van-Nam Dinh[1(✉)], Mau-Viet Ho[1], Van-Cuong Nguyen[1], Tung-Son Ngo[2], and Effiong Charles[3]

[1] Electronic and Communication Faculty, Information and Communication Technology University, Thai Nguyen University, Thai Nguyen, Vietnam
{dvnam,hmviet,nvcuong}@ictu.edu.vn
[2] Computer Fundamental Department, FPT University, Ha Noi, Vietnam
sonnt5@fpt.edu.vn
[3] LIRMM, 161, Rue ADA, 34095 Montpellier Cedex 5, France
talkwithcharles@gmail.com

Abstract. Network-on-Chip (NoC), an interesting paradigm, is one of the newest technologies for VLSI design. In this research, we approach the architecture, algorithms and the performance analyses of a NoC system. The highlight point of this research is implementing additional features to the embedded codes and evaluating special applications in a NoC system based on NOXIM simulator. The results indicated that in certain cases, we should know which is appropriate algorithm for implementing tasks. The evaluating a NoC's performance is an important research trend, hence this study provides one new method for doing that.

Keywords: NoC architectures · NOXIM network-on-chip

1 Introduction

The emergence of Network-On-Chip (NoC) paradigm in recent years has changed the viewing of the world in IP cores design and on-chip communication mechanism. In NoCs, routing determines how information will be transferred among Intellectual Property (IP) cores. Routing algorithms play a vital role in the overall performance and resource utilization of the network, hence the focus of this research work would implement the new method assisting in performance of NoC's platform such some previous researches [8–10].

In this study, we aim to propose a new methodology of analyzing a NoC system based on three main goals. First, we consider the impacts of several NOXIM parameters, then we implement some basic routing algorithm and explore their impact on performance metrics such as delay (cycles), the bandwidth via throughput (flits/IP/cycle) and the power energy (J).

2 Methods

In this section, we detail the main tasks of this study, which includes implementing routing algorithms, techniques, tools and data analysis.

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_64

2.1 Routing Algorithms

Routing determines the path a packet takes from its IP (Source IP core) to the destination IP core. There are two main processes in a routing algorithm: routing function and selection function [3]

Figure 1 presents a 2D-mesh (4×4) topology of Network-on-Chip with all necessary components of a system. A link is used to connect routers together and creates a path for data flow in the network. The links shown in the figure represent both input and output links.

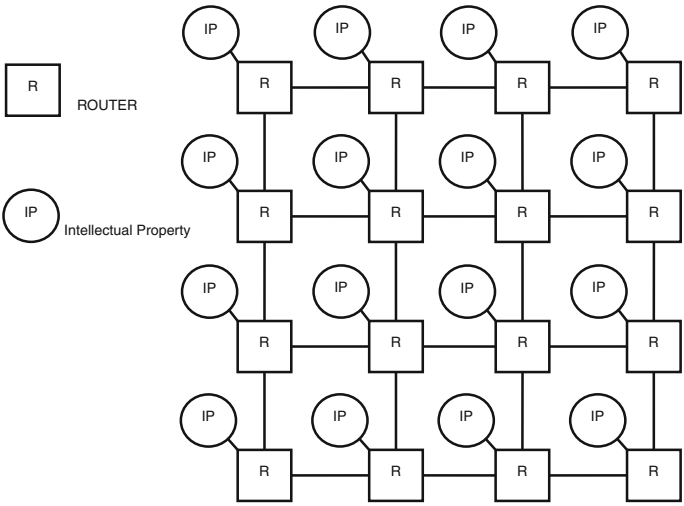


Fig. 1. 2D-mesh topology(4×4)

In this work, we focus our attention solely on the routing algorithm. As mentioned early, the routing algorithm impacts the overall performance of the system in terms of latency and power.

Deterministic Algorithms: Figure 2 depicts the mechanism of a deterministic algorithm in NoC system. In this situation, the selection block is not presented because the routing function returns only one fix path that have already created before transmitting process.

Adaptive Algorithms: In this mechanism, the network status information such as link utilization and buffer's status are cared by selection function block. Figure 3 shows the mechanism of this algorithm. In adaptive routing, several paths from source IP core to destination IP cores are available. So, it leads to some disadvantages such as needing large memory and implementation complexity [7]

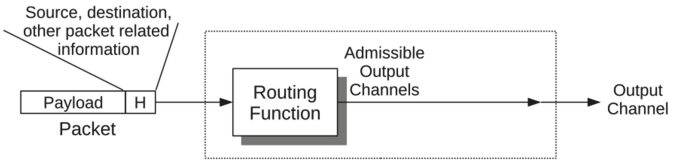


Fig. 2. The deterministic algorithm mechanism [3]

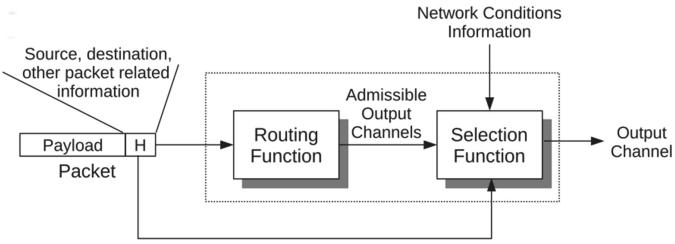


Fig. 3. The mechanism of adaptive algorithms [3]

In this study, we exploited deterministic algorithm because of it's advantages such as suitable with an uniform or regular traffic patterns. The other reason is suitable with a real technology system such as 2D-mesh topology, This leads to available in implementing on a real system.

2.2 A Methodology for Data Analyzing

Figure 4 shows the data flow of our methodology. Each of the process will be discussed in some detail in the following subsection.

In the methodology of implementing, at first, we choose applications [7] for inputs parameters configuring. Then, system will be configured with these basic information before creating proposal scenarios to implement based on noxim platform support [2] for embedded codes [6].

Te evaluate the system performance via Noxim, we implement this method with some specific applications that will be used for configuration the system's parameters. Specially, for updating the information to route packet during communicating processes between source to destination IP cores.

System's Inputs: We use the applications shown in Fig. 5 as our test cases. They are DVOPD, MPEG-4 and MWD applications [4].

System's Configuration: This research includes some scripts that are very useful because it allows the designers can integrate new tools such as ShellScript and Gnuplot for performance evaluation simulations [6].

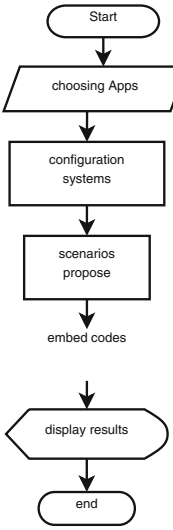


Fig. 4. The data flow of methodology.

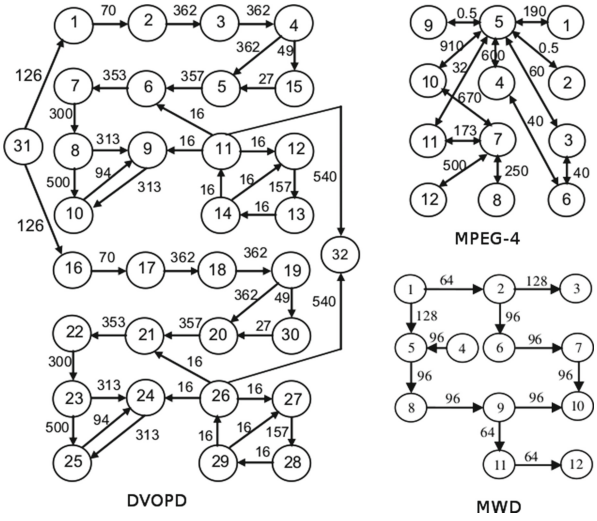


Fig. 5. The system's input applications [4].

In this study, we contributed one new method using NOXIM to map application into system, then with each routing algorithm in Network-on-Chip system (see on Fig. 6)

Futures	Description
Network configuration	
Topology	8x8 MESH 2D
Control Flow	Credit Based Mechanism
Routing algorithm	Deterministic XY algorithm
Switch technique	Wormhole (WH) switching
Communication pattern	
Traffic pattern	Complement, uniform
Packet size	2, 4, ... 512,1024 flits
Simulation environment	
HDL	SystemC 2.3.1
Operating system	Linux (Ubuntu 15.10)

Fig. 6. The system configuration

Proposed Evaluation Scenarios: This section will be used for simulating and collecting the final results that aim at exploring how different simulation parameters affects the system performance.

The main idea is that for each routing algorithm, we evaluate the impact of different packet injection rates for the applications.

In each scenario, the inputs consist of the core application graph (.cg), the reconfiguration file (.yaml); and the outputs include latency (cycles), throughput (flits/IP/cycle) and the power energy (J).

We considered DVOPD, MPEG-4 and MWD applications for scenario 1, 2 and 3 respectively.

We can convert from Fig. 5 into the core graph as Fig. 7 to make the simulation.

Embedded Codes: For this work, we integrated GNUplot tool (the information in Fig. 8) and techniques for system simulation such as ShellScript (open source from link [6]) into NOXIM.

2.3 System's Performance

As mentioned before, we used the data transferring latency (cycles), bandwidth of channels between IP cores, and power/energy (J) are main parameters for estimating the NoC system's performance.

```
% Communication graph of DVOPD
%       Source  targe
        1       2
        2       3
        3       4
        4       5
        4       15
        5       6
        6       7
        7       8
        8       9
        8       10
        9       10
        10      9
        11      9
        11      6
        11      12
        11      32
        12      13
        13      14
        14      11
        14      12
        15      5
        16      17
        17      18
        18      19
        19      20
% Communication graph of MPEG-4 application
%       Source  targe
        1       5
        5       1
        2       5
        5       2
        5       9
        9       5
        10      5
        5       10
        4       5
        5       4
        3       5
% Communication graph of MWD
%       Source  targe
        1       2
        2       3
        2       6
        1       5
        4       5
        6       7
        7       10
        5       8
        8       9
        9       10
        9       11
        11      12
```

Fig. 7. Communication graphs data inputs.

```
G N U P L O T
Version 4.6 patchlevel 6    last modified September 2014
Build System: Linux i686

Copyright (C) 1986-1993, 1998, 2004, 2007-2014
Thomas Williams, Colin Kelley and many others
```

Fig. 8. GNUplot tool.

Latency is one of the most important parameter to estimate the performance of system. It is the time difference (in clock cycles) between when a packet gets delivered to its destination IP and when it was sent.

The latency of the system can be follow the formula (B.1) [7]

$$L_{avg} = \frac{1}{1} \sum \frac{1}{N_i} \sum L_{ij} \tag{B.1}$$

Throughput is the parameter that was used for evaluating the bandwidth of data transfer between IP cores. It can be used to estimate the quality of service the network can achieve.

The throughput of the system can be follow the formula (B.2) [7]

$$T_{avg} = \frac{1}{N(T_{sim} - T_{warm})} \sum N_i \tag{B.2}$$

Power/Energy Consumption: is also a vital parameter that needs to be taken into consideration especially in the area of developing an integrated circuits.

3 Results

For estimating of the final results in this study, we will focus on the system performances such as latency, which is the time between the emission of data and reception of this data at the destination (cycles); throughput, which defines the quality of data transmitted from source to destination per a time unit (flits/IP/cycle) [5] and Power Energy (J).

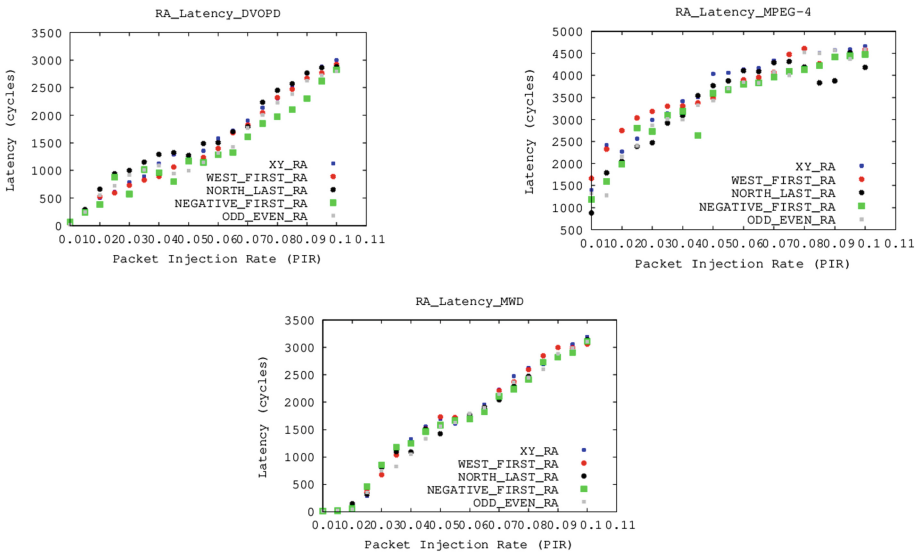


Fig. 9. The latency (cycles) of the system.

Figures 9 to 11 depict the outputs in terms of latency, throughput and power/energy corresponding to each application. First, the applications were mapped onto the NoC and NOXIM was used to run simulations with low to high packet injection rate (PIR).

The presented results provides designers with key insight into how different mappings and routing impacts NoC performance for a given application.

This will enable designers choose suitable routing algorithms for applications to be mapped on NoC. Our results show that the choice of suitable routing algorithm depends on the application requirements. If the system has real time constraint, as in Fig. 9, The NEGATIVE_FIRST algorithm is suitable. This is the case for DVOPD application. However, for MPEG-4 application, with bandwidth as highest priority factor, Fig. 10 suggest that the NORTH_LAST algorithm is most suitable.

The results obtained in this work is based on reconfiguring of the NOXIM platform [1], and we believe that similar evaluation will be obtained on a fabricated NoC system given our constraints.

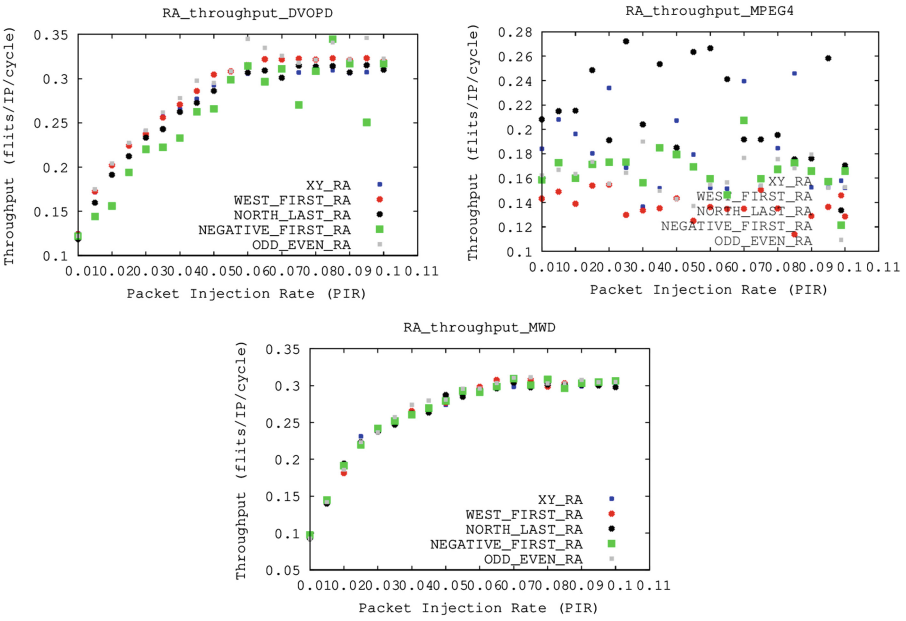


Fig. 10. The throughput (Flits/cycle/IP) of the system.

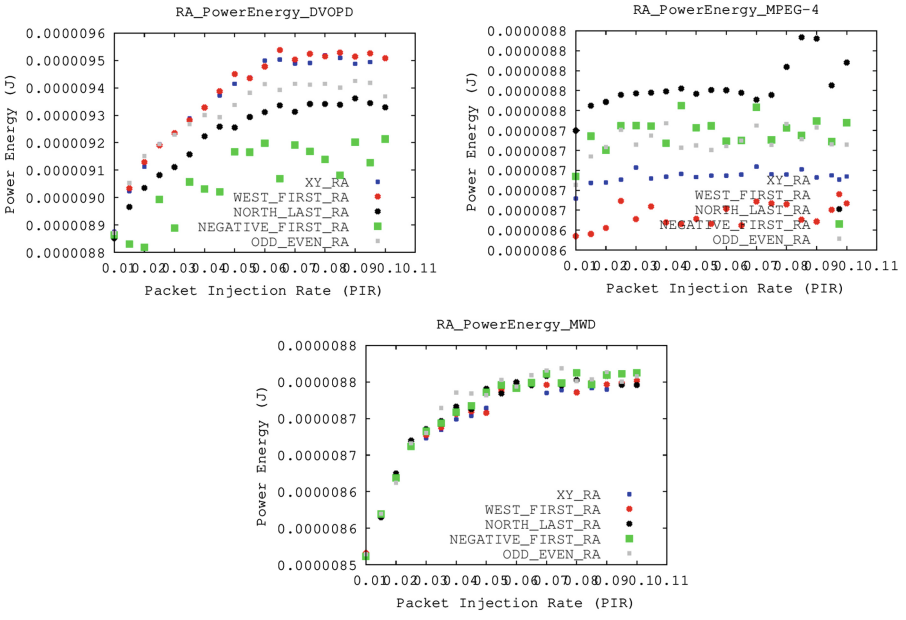


Fig. 11. The Power energy (J) of the system.

Acknowledgment. The authors would like to thank to the research group of University of Catana (Italy), especially Professor Maurizio Palesi, who helped and gave us a lot guides during we have doing this study.

We also would like to thank to Professor Xuan-Tu TRAN, who gave many useful advice for discussing section and also helping us can complete this paper with the best ability in his revising.

References

1. Catania, V., Mineo, A., Monteleone, S., Palesi, M., Patti, D.: (Italy) Noxim: an open, extensible and cycle-accurate network-on-chip simulator. In: 2015 IEEE International Conference on Application-Specific System, Architectures and Processors, pp. 27–29, Toronto, Canada, July 2015
2. NOXIM: Open Source. https://github.com/davidepatti/noxim
3. Palesi, M., Daneshtalab, M.: Routing Algorithms in Network-On-Chip. Springer Science + Business Media, New York (2014)
4. Chattopadhyay, S.: A survey on application mapping strategies for network-on-chip design. J. Syst. Archit. **59**, 60–75 (2013)
5. Xuan-Tu, T., et al.: Emerging Aspects in Electronic and Communication Engineering, pp. 91–92. Vietnam National University Publisher, Hanoi (2013)
6. ICTA2016: Open source. https://github.com/vnd2586/ICTA2016
7. Singh, J.K., et al: Performance evaluation of different routing algorithms in network on chip. In: 2013 IEEE Asian Pacific Conference on Postgraduate Research in Microelectronics and Electronics (2013)
8. Le Van, T.V., Tran, X.T.: Simulation and performance evaluation of a network-on-chip architecture based on systemC. In: Proceedings of the 5th International Conference on Advanced Technologies for Communication (ATC2012), pp. 170–175, Hanoi, October 2012
9. Dang, N.K., Le Van, T.V., Tran, X.T.: FPGA implementation of a low latency and high throughput network-on-chip router architecture. In: Proceedings of the 2011 International Conference on Integrated Circuits and Devices in Vietnam (ICDV 2011), Hanoi, August 2011
10. Le Van, T.V., Ngo, D.T., Tran, X.T.: A SystemC based simulation platform for network-on-chip architecture. In: Proceedings of the 2012 IEICE International Conference on Integrated Circuits and Devices in Vietnam (ICDV 2012), pp. 132–136, Danang, August 2012

The Asynchronous Cooperative Amplify-and-Forward Relay Network with Partial Feedback to Improve the System Performance

The-Nghiep Tran[1(✉)], Van-Bien Pham[2], and Huu-Minh Nguyen[1]

[1] Faculty of Radio-Electronics, Le Quy Don Technical University, Hanoi, Vietnam
nghiepsqtt@gmail.com, huuminhtsqtt@gmail.com
[2] Vietnam People's Army Broadcast Center, Hanoi, Vietnam
http://www.lqdtu.edu.vn

Abstract. In this paper, a new near-optimum detection (NOD) scheme is combined with partial feedback technique for two dual-antenna relay nodes in the amplify-and-forward (AF) asynchronous cooperative relay network. The application of partial feedback in the proposed scheme not only offers to reduce a cooperative relaying process due to not using distributed close loop extended orthogonal space-time block code (DCL EO-STBC) encoder, but also improves the end-to-end signal noise ratio (SNR). Moreover, a near-optimum detection (NOD) scheme is used at the destination to remove completely the interference components induced by inter-symbol interference (ISI) among the relay nodes. The analysis and simulation results demonstrate that the performance of the proposed scheme outperforms the previous feedback scheme in both the perfect synchronization and imperfect synchronization assumption cases.

Keywords: Partial feedback · Near-optimum detection · Inter-symbol interference · Distributed close loop extended orthogonal space-time block code · Asynchronous cooperative relay

1 Introduction

Communicating via multiple cooperative relay nodes helps to combat the detrimental effects of fading in wireless channels by providing additional diversity/reliability to the system. This diversity arises because the channel fading gain varies depending on the location of the relay node and thus well separated multiple relay nodes can bring in independent channel fading gains and hence more diversity or protection to the transmitted information [1]. However, the assumption that the cooperative relay nodes transmit the corresponding symbols to the destination node at the same time in a perfect synchronized manner is difficult or impossible. Thus, the bit error rate (BER) performance significantly degrades due to the inter-symbol interference (ISI) in the cooperative relay transmissions [2–8].

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_65

The most commonly utilized strategies with various cooperative relaying protocols are either amplify and forward (AF) or decode and forward (DF) [9]. In [2] and [4], asynchronous cooperative relay networks are demonstrated using distributed close loop extended orthogonal space-time block code (DCL EO-STBC) and distributed space-time block code (D-STBC) with near-optimum detection (NOD) at the destination, respectively. These proposed schemes overcame this lack of the imperfect synchronization by introducing ISI components with DF strategy, which leads to high process at relay nodes. Another example of the DF protocol, in [8], Elazreg et al. propose using distributed close loop quasi orthogonal space-time block code (DCL QO-STBC) and Sub-optimum detection. Whereas in [1] and [3] authors considered a simple cooperative relaying strategy with DCL EO-STBC and DCL QO-STBC, respectively. In this cooperative method, each relay node receives a noisy version of the signal transmitted by the source node, and amplifies and retransmits this noisy version to the destination node. The transmission strategy calls AF protocol, which does not need channel information at the relay nodes, but full channel information from the source node to the relay nodes and from the relay nodes to the destination node are required at the destination node. In addition, the AF strategy does not perform any decoding operation at the relay node. Therefore, the AF protocol is a best cooperative relaying process in the case of relay nodes require low operation and saving power such as Ad hoc wireless network or sensor network [10].

All above designs can effectively eliminate ISI components with AF or DF protocol and exploit the diversity gain and transmission rate. But, the process complexity of relay node is still high because the relay nodes need to use a DCL EO-STBC (DCL QO-STBC) encoder.

This paper proposes a new AF asynchronous cooperative relay network with partial feedback. The AF strategy of proposed scheme reduces a cooperative relaying process due to not using DCL EO-STBC encoder in comparison with the previous cooperative AF relay network [1]. Moreover, the proposed scheme not only offers to achieve the maximum signal noise rate (SNR), but also can completely eliminate ISI components with implementing near-optimum detection at the destination.

The remainder of this paper is organized as follows: In the Sect. 2, the proposed AF asynchronous cooperative relay network with partial feedback is described; The detection of interference cancellation is presented and analyzed in the Sect. 3; Simulation results and performance comparisons are presented in Sect. 4; Finally, Sect. 5 summarizes this paper.

In the remaining part of this paper, $[.]^T$, $[.]^*$ and $\|.\|^2$ denote transpose, complex conjugate, and Frobenius operation, respectively; $\Re$ and $\Im$ present to take the real and imaginary part of the complex variable, respectively; $\mathbb{E}[.]$ represents an expectational operation; and $\mathcal{A}$ indicates the signal constellation.

2 The Proposed AF Asynchronous Cooperative Relay Network with Partial Feedback

In this paper, the AF asynchronous cooperative relay network with partial feedback is considered as shown in Fig. 1. This model consists of the single-antenna source node and destination node, and two dual-antenna relay nodes. The direct transmission link from the source node to the destination node is ignored due to the effect of path loss and the limited transmitted power. Let f_{ik} denotes the channel coefficient from the source node to i-th the antenna of the k-th relay node and g_{ik} is the channel coefficient from the i-th antenna of the k-th relay node to the destination node. We also assume that all channel coefficients f_{ik} and g_{ik} (for $i,\ k = \overline{1,2}$) are kept constant during two symbol interval and varied randomly in the next two symbol interval (i.e. a flat-static channel). The noise terms of both relay and destination node are assumed the additive complex Gaussian noises with the distribution $\mathcal{CN}(0,1)$. It is assumed that feedback link has no feedback error and no delay. The total transmitted power utilized for the transmission of one symbol is fixed as P(dB). We adopt the optimal power allocation presented in [11] as follows:

$$P_1 = \frac{P}{2},\ P_2 = \frac{P}{4}, \tag{B.1}$$

where, P_1, P_2 are the transmit power at the source and the each relay node, respectively.

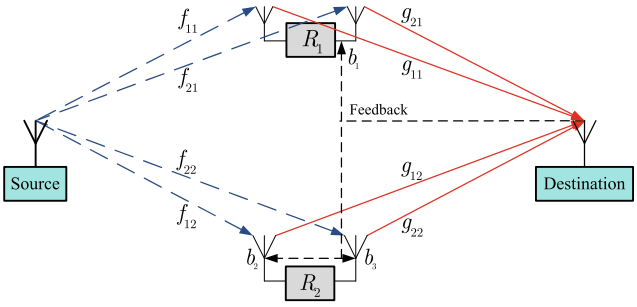


Fig. 1. The AF asynchronous cooperative relay network with partial feedback.

In the first phase, the source node broadcasts the sequent modulated symbols $s(n)$ to every relay node in one symbol period. So, the received symbol at the i-th antenna of the k-th relay node is presented as:

$$r_{ik}(n) = \sqrt{P_1} f_{ik} s(n) + z_{ik}(n), \tag{B.2}$$

where, $z_{ik}(n)$ is noise at the i-th antenna of the k-th relay node.

In the second phase, this paper applies a partial feedback in [12] to achieve the array gain for the AF asynchronous cooperative relay network. It is assumed

that b_1, b_2 and b_3 are the feedback bits from the destination node to the relay nodes. Firstly, the received symbols in (B.2) are multiplied with feedback bits which is equivalent to phase rotations of the corresponding channel coefficients. This is important to improve end-to-end SNR of the proposed scheme. The transmitted symbol vector for two dual-antenna relays is presented as in (B.3):

$$E_B = [r_{11}(n)\ b_1 r_{21}(n)\ b_2 r_{21}(n)\ b_3 r_{22}(n)] \tag{B.3}$$

where, b_i $(i = 1, 2, 3)$ is equal to 1 or -1 depending on the channel static information.

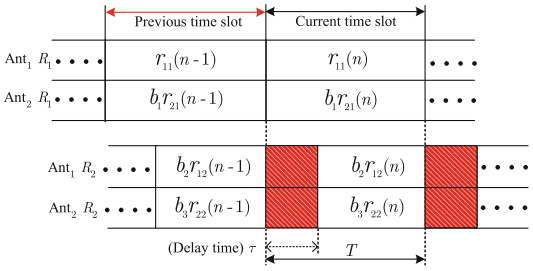


Fig. 2. The effect of imperfect synchronization in the asynchronous cooperative relay network.

As the previous discussions, the propagation delays of received signals from the distinct relay node are not the same length due to the different distances between the relay nodes and the destination node. Without loss of generality, it is assumed that the links of the between both antennas of the first relay node (denotes R_1) and the destination are synchronized perfectly. But, the links of the between both antennas of the second relay node (denotes R_2) and the destination are synchronized imperfectly. So that is $\tau \neq 0$ as shown in Fig. 2. The received signal at the destination node are presented as following:

$$\begin{aligned} y(n) = &\sqrt{\frac{P_2}{(P_1+1)}}\left[r_{11}(n)g_{11} + b_1 r_{21}(n)g_{21} + b_2 r_{12}(n)g_{12} + b_3 r_{22}(n)g_{22}\right] \\ &+ \sqrt{\frac{P_2}{(P_1+1)}}\left[b_2 r_{12}(n-1)g_{12}(-1) + b_3 r_{22}(n-1)g_{22}(-1)\right] + v(n) \end{aligned} \tag{B.4}$$

$$= \sqrt{\frac{P_1 P_2}{(P_1+1)}} h s(n) + I_{\text{int}}(n) + w(n), \tag{B.5}$$

where, $h = f_{11}g_{11} + b_1 f_{21}g_{21} + b_2 f_{12}g_{12} + b_3 f_{22}g_{22}$ is the equivalent channel gain. $I_{\text{int}}(n)$ is the interference term from both antennas of the second relay node due to the effect of asynchronous phenomenon and calculates as following:

$$I_{\text{int}}(n) = \sqrt{\frac{P_1 P_2}{(P_1+1)}}\left\{b_2 f_{11}g_{12}(-1) + b_3 f_{22}g_{22}(-1)\right\} s(n-1). \tag{B.6}$$

$w(n)$ is the equivalent noise at destination and it is presented as:

$$w(n) = \sqrt{\frac{P_2}{(P_1+1)}}\left(g_{11}z_{11}(n) + b_1 g_{21} z_{21}(n) + + b_2 g_{12} z_{12}(n) + b_3 g_{22} z_{22}(n)\right) + v(n). \quad \text{(B.7)}$$

$v(n)$ is noise at the destination node. As shown in Fig. 2, $g_{i2}(-1)$ $(i = 1, 2)$ is the channel gain between the antenna of R_2 and the destination under imperfect synchronization. The $g_{i2}(-1)$ is related with g_{i2} as following [1]:

$$\beta = |g_{i2}(-1)|^2 / |g_{i2}|^2; \quad i = 1, 2 \quad \text{(B.8)}$$

This paper adopted with $\beta = 0$ for $\tau = 0$ and $\beta = 1$ (i.e. 0 dB) for $\tau = 0.5T$ [1]. Note that, the factor $\sqrt{P_2/(P_1+1)}$ in the Eq. (B.5) is ensured the average transmission power at the each relay node is P_2. It is assumed that the destination node knows completely all the channel coefficients f_{ik} and g_{ik}. Its knowledge of the channels can be obtained by sending training signals. Then, the conventional detection procedure of the proposed scheme can be carried out via the Least Square (LS) method as follows:

$$y(n) = h^* r(n) = \sqrt{\frac{P_1 P_2}{(P_1+1)}}\lambda s(n) + h^* I_{\text{int}}(n) + h^* w(n) \quad \text{(B.9)}$$

$$\tilde{s}(n) = \arg\min_{s_m \in \mathcal{A}} \left| y(n) - \sqrt{\frac{P_1 P_2}{(P_1+1)}}\lambda s_m \right|^2, \quad \text{(B.10)}$$

where, $\mathcal{A}$ denotes the constellation and $\lambda = h^* h = \alpha_B + \beta_B$ is total performance gain as following:

$$\alpha_B = |f_{11}g_{11}|^2 + |f_{21}g_{21}|^2 + |f_{12}g_{12}|^2 + |f_{22}g_{22}|^2; \quad \text{(B.11)}$$

$$\begin{aligned}\beta_B = {} & 2b_1 \Re\left(f_{11}g_{11}f_{21}^* g_{21}^*\right) + 2b_2 \Re\left(f_{11}g_{11}f_{12}^* g_{12}^* + b_1 f_{21}g_{21}f_{12}^* g_{12}^*\right) \\ & + 2b_3 \Re\left(f_{11}g_{11}f_{22}^* g_{22}^* + b_1 f_{21}g_{21}f_{22}^* g_{22}^* + b_2 f_{12}g_{12}f_{22}^* g_{22}^*\right).\end{aligned} \quad \text{(B.12)}$$

The proposed scheme gets the conventional diversity gain α_B and the array performance gain β_B. It is clear that $\alpha_B \geq 0$ is always true. The application of the partial feedback idea in [12], the additional array gain can achieved if the feedback bits are calculated as following Table 1.

Remarks

- The additional array gain of the proposed scheme β_B in (B.12) is bigger than the performance array gain of the DCL EO-STBC scheme λ_f in [1] when the partial phase feedback bits are calculated as Table 1. In consequence, the signal noise ratio (SNR) of the proposed scheme is better than the DCL EO-STBC scheme.
- It is clear that the process of relay nodes is very simple due to relay nodes do not use neither the DF protocol or the DCL EO-STBC encoder.
- The Eq. (B.9) shows that the performance of new scheme will damage because there is an ISI component $I_{\text{int}}(n)$ in the received signal unless the $g_{i2}(-1)$ gets 0 (i.e. $\tau = 0$, the case of perfect synchronization) or the ISI component is removed.

Table 1. Calculation of the feedback bits.

Step 1: b_1	$= 1 \quad if\ \Re\left(f_{11}g_{11}f_{21}^{*}g_{21}^{*}\right) \geq 0$
	$= -1 \quad if\ \Re\left(f_{11}g_{11}f_{21}^{*}g_{21}^{*}\right) < 0$
Step 2: b_2	$= 1 \quad if\ \Re\left(f_{11}g_{11}f_{12}^{*}g_{12}^{*} + b_1 f_{21}g_{21}f_{12}^{*}g_{12}^{*}\right) \geq 0$
	$= -1 \quad if\ \Re\left(f_{11}g_{11}f_{12}^{*}g_{12}^{*} + b_1 f_{21}g_{21}f_{12}^{*}g_{12}^{*}\right) < 0$
Step 3: b_3	$= 1 \quad if\ \Re\left(f_{11}g_{11}f_{22}^{*}g_{22}^{*} + b_1 f_{21}g_{21}f_{22}^{*}g_{22}^{*} + b_2 f_{12}g_{12}f_{22}^{*}g_{22}^{*}\right) \geq 0$
	$= -1 \quad if\ \Re\left(f_{11}g_{11}f_{22}^{*}g_{22}^{*} + b_1 f_{21}g_{21}f_{22}^{*}g_{22}^{*} + b_2 f_{12}g_{12}f_{22}^{*}g_{22}^{*}\right) < 0$

3 Near-Optimum Detection for the Proposed Scheme

As the previous mention in Sect. 2, the key to mitigate the above impact of imperfect synchronization is to remove the ISI component of $I_{\text{int}}(n)$ in (B.5). Fortunately, the Fig. 2 and Eq. (B.6) show that there are only two previous ISI components in received signals at the destination. Therefore, the NOD scheme is used at the destination to remove completely the ISI components $I_{\text{int}}(n)$ in the Eq. (B.5) before applying the matched filter in (B.9). In fact, $s(n-1)$ is already known if the detection process has been initialized properly [1]. So, the interference components $I_{\text{int}}(n) = \sqrt{P_1P_2/(P_1+1)}\left\{b_2 f_{11}g_{12}(-1) + b_3 f_{22}g_{22}(-1)\right\}s(n-1)$ in (B.5) can be eliminated completely as the below procedure:

Step 1: Remove the ISI component $I_{\text{int}}(n)$ in (B.5)

$$y'(n) = y(n) - I_{\text{int}}(n) = \sqrt{\frac{P_1P_2}{(P_1+1)}}hs(n)+w(n), \tag{B.13}$$

Step 2: Linear transform

$$y''(n) = h^{*}y'(n) = \sqrt{\frac{P_1P_2}{(P_1+1)}}\lambda s(n)+h^{*}w(n), \tag{B.14}$$

Step 3: Application the LS (Least Square):

$$\tilde{s}(n) = \arg\min_{s_m \in \mathcal{A}} \left| y''(n) - \sqrt{\frac{P_1P_2}{(P_1+1)}}\lambda s_m \right|^2. \tag{B.15}$$

As above analysis, the ISI component in the Eq. (B.13) is removed if the initialized signal $s(n-1)$ has no decision feedback error, then the above procedure is optimum in terms of maximum likelihood which calls Near-Optimum Detection (NOD).

Remarks

– It is clear that the above procedure can remove completely the ISI component which cause the different propagation delays in the asynchronous cooperative relay network.

– The above procedure does not rely on the detection results of the transmission link between the source node and destination, so the detection complexity of the proposed scheme is significantly reduced as compared with Sub-optimum detection [8].

4 Simulation Results

In this section, simulation results with the Monte-Carlo method are shown and demonstrated for our previous analysis. All schemes use QPSK modulation and have the same total transmitted power. The fading is assumed quasi-static fading which keeps a constant within a frame and changes independently from frame-to-frame. The level of imperfect synchronization (i.e. τ the time delays between both antennas of the second relay node and the destination node) is changed by varying the value of β (dB).

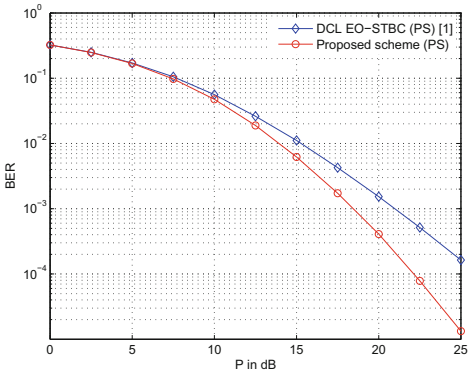
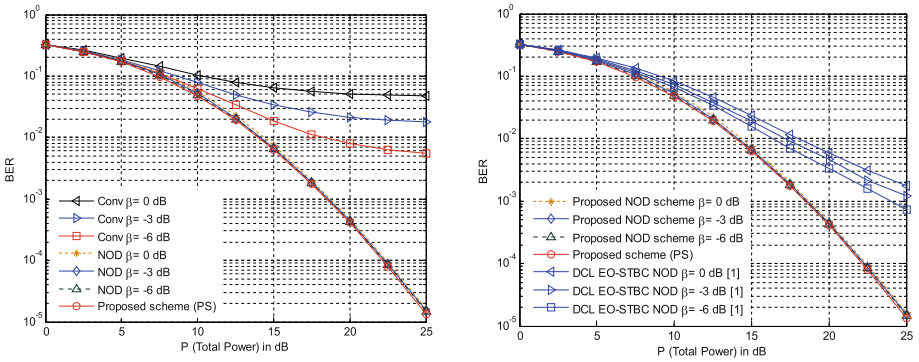


Fig. 3. BER performance comparison of proposed scheme and DCL EO-STBC scheme [1] in case perfect synchronization.

Figure 3 shows the comparison results of BER performance the proposed scheme and DCL EO-STBC scheme [1] in case perfect synchronization (PS). It confirms that the proposed scheme is better than DCL EO-STBC scheme because the SNR of the proposed scheme is improved when the simple partial feedback is applied as analysis in Sect. 2. For example, the proposed scheme just requires approximately the source power 18.5 dB to achieve a BER of 10^{-3}, while the DCL EO-STBC scheme requires 21 dB. It notices that two schemes have same hardware requirement.

The BER performances of the proposed scheme utilized NOD or conventional (Conv) detection show in Fig. 4(a) under the different levels of imperfect synchronization ($\beta = 0,\ -3\ and\ -6$ (dB)). Figure 4(a) confirms that the proposed scheme with application of NOD has good performance because the ISI components can completely remove before using the LS detector while Conv detector



a) NOD detector vs Conv detector.

b) The proposed scheme vs the DCL EO-STBC scheme [1].

Fig. 4. The comparison of BER performance proposed scheme.

fails even under small β values (or slight imperfect synchronization). Specially, the BER total system performance of the proposed NOD scheme can be close its BER in the perfect synchronization case. Therefore, the new NOD scheme can fight extremely asynchronous phenomenon.

Comparisons of the BER performance between the proposed scheme and the DCL EO-STBC scheme [1] with the same configuration model and NOD detector are shown in Fig. 4(b). As seen in the simulation results, it is clear that the degradation of BER performance of DCL EO-STBC scheme is very fast when increasing factor β (i.e. more loss of synchronization). Whereas BER performance of proposed scheme degrades very slowly in the asynchronous channel conditions.

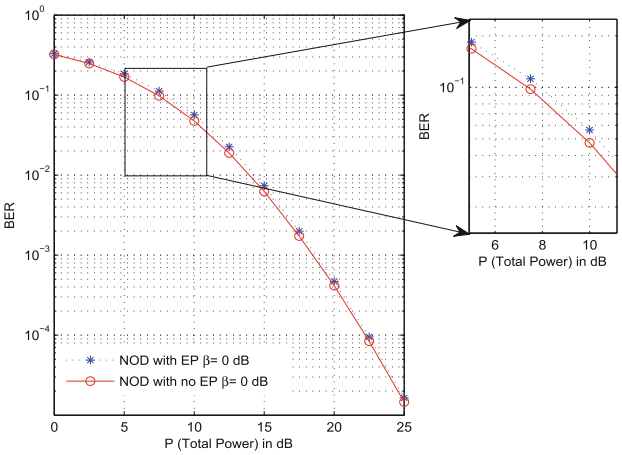


Fig. 5. The impact of error propagation (EP) on the BER performance.

To examine the effect of decision feedback errors, BER performance of the proposed NOD scheme is carried out (i) with error propagation (EP, i.e. $s(n-1)$ gets the value of previous detection or natural propagation of any feedback errors), and (ii) with no EP (i.e. using the previous true symbol) for ISI removal. The BER performance of comparison shows as Fig. 5 with large level of imperfect synchronization $\beta = 0$ dB. The Fig. 5 shows that the impact of error propagation of the near-optimum detector is very minor.

5 Conclusion

In this paper, we have proposed a new AF asynchronous cooperative relay network with partial feedback. The proposed scheme achieves maximum SNR by using the partial feedback while reduces the AF cooperative relaying process without utilized DCL EO-STBC encoder at relay nodes. Analysis and simulation results confirmed that the application of NOD at destination can completely remove the ISI components and overcome this lack of asynchronous cooperative relay network. In practice, the 3 feedback bits of the new scheme requirement are available in the 3G system with supporting the 1.5 kbps feedback link. Therefore, this scheme is fully compatible with cooperative relay networks such as wireless ad hoc or sensor networks in asynchronous channel conditions.

References

1. Qaja, W., Elazreg, A., Chambers, J.: Near-optimum detection for use in closed-loop distributed space time coding with asynchronous transmission and selection of two dual-antenna relays. In: Wireless Conference, Guildford, UK, pp. 1–6 (2013)
2. Qaja, W.M., Elazreg, A.M., Chambers, J.A.: Near-optimum detection scheme with relay selection technique for asynchronous cooperative relay networks. IET Commun. **8**(8), 1347–1354 (2014)
3. Qaja, W.M., Elazreg, A.M., Chambers, J.A.: Distributed space time transmission with two relay selection and parallel interference cancellation detection to mitigate asynchronism. In: Proceedings of the Computer Modeling and Simulation (EMS), Valetta, Malta, pp. 220–225 (2012)
4. Zheng, F.C., Burr, A.G., Olafsson, S.: Near-optimum detection for distributed space-time block coding under imperfect synchronization. IEEE Trans. Commun. **56**(11), 1795–1799 (2008)
5. Astal, M., Olivier, J.C.: Distributed closed-loop extended orthogonal STBC: improved performance in imperfect synchronization. In: Proceedings of the Personal Indoor and Mobile Radio Communications, London, England, pp. 1941–1945 (2013)
6. Zheng, F.C., Burr, A.G., Olafsson, S.: Distributed space-time block coding for 3 and 4 relay nodes: imperfect synchronisation and a solution. In: Proceedings of the International Symposium on Personal, Indoor and Mobile Radio Communications, Athens, Greece, pp. 1–5 (2007)
7. Elazreg, A.M., Chambers, J.A.: Sub-optimum detection scheme for asynchronous cooperative relay networks. IET Commun. **5**(15), 2250–2255 (2011)

8. Elazreg, A., Kharaz, A.: Sub-optimum detection scheme for distributed closed-loop quasi orthogonal space time block coding in asynchronous cooperative two dual-antenna relay networks. In: Mumtaz, S., Rodriguez, J., Katz, M., Wang, C., Nascimento, A. (eds.) WICON 2014. LNICST, vol. 146, pp. 217–228. Springer, Heidelberg (2015). doi:10.1007/978-3-319-18802-7_30
9. Nosratinia, A., Hunter, T.E., Hedayat, A.: Cooperative communication in wireless networks. IEEE Commun. Mag. **42**(10), 74–80 (2004)
10. Murthy, C.S.R., Manoj, B.S.: Ad Hoc Wireless Networks Architectures and Protocols. Prentice Hall PTR, Upper Saddle River (2004)
11. Yindi, J.: Combination of MRC and distributed space-time coding in networks with multiple-antenna relays. IEEE Trans. Wirel. Commun. **9**(8), 2550–2559 (2010)
12. Van Pham, B., Pham Truong, S., Pham Hong, T.: Enhancing performance of four transmit antennas and one receive antenna system with partial feedback. In: Proceedings of the Advanced Technologies for Communications (ATC), Hanoi, Vietnam, pp. 259–262 (2012)

The Optimization Model for Calculating Distribution System Planning Integrated Photovoltaic

V.V. Thang(✉)

Thainguyen University of Technology (TNUT), Thai Nguyen, Vietnam
thangvvhtd@tnut.edu.vn

Abstract. A two-stage model for the optimal planning of distribution systems with the presence of photovoltaic generation system (PV) is presented in this paper. The proposed model can determine the optimal sizing and time-frame of the equipment (feeders and transformer substations) in distribution systems. Therefore, the optimal displacement, sizing, technology and installation period of PV are also determined. The objective function is the life cycle costs minimizing of the planning project. The technical constraints are used to guarantee the operability of the distribution system including AC power flow, feeder and substation upgrading section, limited of nodal voltage and PV capacity. The binary variables are also employed in the model to represent the cost function of the equipment as well as the investment and upgrade decisions. The algorithm is programmed in GAMS. The feasibility and effectiveness of the proposed model are examined in a test system.

Keywords: Distribution systems planning · Photovoltaic

1 Introduction

In the past decade, the planning of distribution systems had a major change due to the implementation of competitive electricity markets (EM), distributed generator (DG) technology and environmental pollution issues. In particular, DG which is directly connected to DS or directly supplying customers is used widely. These sources may employ various primary-mover technologies such as gas turbines, combined heat and power, fuel cells, wind turbines and photovoltaic. Therefore, the benefits of DG can be thought as the reduction of transmission and distribution cost, power loss and the enhancement of flexibility and reliability of DS, the improvement of differential voltage at nodes, as well as the reduction of environmental pollutions [1]. One of the DG technologies the most interested is PV and has been presented in many previous researches [2,3]. However, PV requires high investments, increases the complexity of measurements and relay protection and the operation of DS [4]. In addition, PV uses renewable energy resources has power generation varying according to natural conditions. Many planning models of the DG with different primary-mover technologies and

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_66

PV in distribution systems have been researched and proposed. The authors in [5] presented a long-term DS planning model in order to determine the capacity, location for a new building investment process or upgrading the existing equipment by using the popular mathematical programming. The objective of the model is minimizing the total investment and operation costs of DG, the investment cost for feeder and substation transformers during the planning period. Another model in [6,7] was proposed with the objective function including the total investment and operation costs of DG, feeders and substation transformers upgrading costs, energy expenses and load's interruption costs. The objective function of the two-stage DS planning model in [8] is the minimization of total costs for upgrading feeders, substation transformers and DG construction, energy expenses purchased from markets and environmental pollution costs. Similarly, Ref. [9] introduced a DS planning model for determining the optimal equipment sizing and time-frame of DS. Besides, the model with the objective function of minimizing the life cycle cost for the distribution system planning was introduced in 10. The model aims to find the best distribution system planning scheme to maximize the overall benefits and costs in the life cycle of the system. In previous studies, the output power of PV is always assumed to be constant without regarding to the natural variability which depends on the primary energy - this is of course far from reality. The power flow constraint usually uses DC model so the impacts of reactive power to the planning problem is ignored. Therefore, this paper proposes a DS optimal planning model that integrates the power output characteristics of PV, characteristics of load demand and electricity price. The AC power flow model is used to consider the influences of reactive power in DS planning. The detail about the planning model is illustrated in the following sections.

2 The Mathematical Model

In competitive EM, DS are usually managed by distribution companies. These companies can buy electricity from wholesale markets or combine with the on-site PV in order to meet load demands. Therefore, the economic and technical measures of planning projects are changed which affects considerably to the duration, upgrading capacity of feeders and substations when PV are chosen in DS. The proposed DS planning model is also executed in two stages. A MINLP model in the first stage is calculated and its results are fed into the second stage to obtain a comprehensive plan. The second stage receives information transferred from the first stage includes a set of decisions on location and period for equipment investment. Therefore, this model needs not use for binary variables and it is a NLP model. The accuracy of the model is added in the second stage to more closely reflect the required investments and production schedules.

2.1 The Mathematical Model of the First Stage

2.1.1 Objective Function

The objective function of proposed model is to minimize the total life cycle cost of the investment project during calculation period as shown in [9] and is added with the part of reduced costs due to lower emissions. The total cost is calculated at the base year by equation with discount rate r as Eq. (1) with the symbols of model are presented in Table 1.

$$J = Min \sum_{t=1}^{T} \frac{1}{(1+r)^t}.(CF_t + CS_t + CPV_t + EPV_t + ES_t - TCO_t + RN_t) \forall t \in T \tag{1}$$

where, the component is the upgrading costs of feeders for year t with fixed capital cost (C^{FF}) and the variable capital cost (C^{FC}) as shown in Eq. (2). The $\alpha_{ij.t}$ is the binary variable to represent the cost characteristic and decision variable for feeder upgrading is $F_{ij.t}$.

$$CF_t = \sum_{i=1}^{N} \sum_{j=i+1}^{N} L_{ij}(C^{FF}.\alpha_{\mathrm{ij}.t} + C^{FC}.F_{ij,t}) \quad \forall ij \in N, i \neq j, t \in T \tag{2}$$

Similar to the above, the substation transformers upgrading costs in year t (CS_t) including fixed capital cost (C_{SF}) and variable capital cost (C_{FC}) is presented in Eq. (3). The $\gamma_{i.t}$ is a binary variable to represent the cost characteristic and decision variable for transformer upgrading is $\Delta.S^S_{i.t}$.

$$CS_t = \sum_{i=1}^{NS} (C^{SF}.\gamma_{i,t} + C^{SC}.\Delta S^S_{i,t}) \qquad \forall i \in N_S, t \in T \tag{3}$$

Electrical energy purchased cost from EM (ES_t) is presented in (4).

$$\mathrm{ES}_t = \sum_{i=1}^{NS} \sum_{s=1}^{SS} \sum_{h=1}^{H} D_s.k_P.(\rho^S_{P.h}.P^s_{i,t,s,h} + \rho^s_{Q.h}.Q^s_{i,t,s,h}) \qquad \forall i \in N_S, t \in T, s \in SS, h \in H \tag{4}$$

The Eq. (5) is new investment costs in year t of PV. Beside, electrical energy purchased cost from EM and costs for operation and maintenance of PV, operation season s and time h are shown in Eq. (6).

$$CPV_t = \sum_{i=1}^{N_{PV}} C^{PV}_i.P^{PV}_{i,t} \quad \forall i \in N_{PV}, t \in T \tag{5}$$

$$EPV_t = \sum_{i=1}^{N_{PV}} \sum_{s=1}^{SS} \sum_{h=1}^{H} D_S.\rho^{PV}_P.P^{PV}_{i,t,s,h} \qquad \forall i \in N_{PV}, t \in T, s \in SS, h \in H \tag{6}$$

Table 1. Variables, sets, indices, parameters and symbols

Symbol	Definition	Symbol	Definition
N	Set of buses in distribution system	i, j	Bus (i, j ∈ N)
N_L	Set of load buses in distribution system	N_S	Set of substation buses in distribution system
N_{PV}	Set of PV buses in distribution system	h, H	Hour and hours per day (h ∈ H)
t, T	Planning year and overall planning period (t ∈ T)	s,SS	Season and total seasons in year (s ∈ SS)
$F_{ij,t}$	Upgrading section of Feeder (mm^2)	$P^{PV}_{i,s,t,h}$	Active output power of PV (kW)
$\Delta S^S_{i,t}$	Addition capacity for Substation (MVA)	$U_{i,s,t,h}$	Voltage for bus (pu)
$P^{PV}_{i,t}$	New investment capacity of PV (MW)	$\delta_{i,s,t,h}$	Voltage angle at bus (pu)
$P^S_{i,s,t,h}$	Active power purchased from electricity market (kW)	$\alpha_{ij,t}$	Binary variable on feeder upgrade decision (1/0)
$Q^S_{i,s,t,h}$	Reactive power purchased from electricity market (kVAr)	$\gamma_{i,t}$	Binary variable on feeder upgrade decision (1/0)
$\Delta S^F_{ij,t}$	Addition capacity of Feeder (MVA)	ρ^{PS}_s	Active power purchased cost from market ($/kWh)
R	Discount rate (%)	ρ^{QS}_s	Reactive power purchased cost from market ($/kVAh)
C^{FF}	Fixed capital cost of Feeder ($/km)	$\rho^{OG}_{P,h}$	O&M cost of PV ($/kWh)
C^{FC}	Variable capital cost of Feeder ($/km.mm²)	$PD_{i,s,t,h}$	Active power demand at bus (kW)
L_{ij}	Length of Feeder (km)	$QD_{i,s,t,h}$	Reactive power demand at bus (kVAr)
$Y_{i,j}, \theta_{i,j}$	Magnitude and Angles of admittance matrix element (pu)	$P^{PV}_{max,i}$	Maximum power limit of PV i (MW)
C^{SF}	Fixed capital cost of Substation ($/Substation)	$P'^{PV}_{max,i}$	New maximum power limit of PV in second stage (MW)
C^{SC}	Variable capital cost of Substation ($/MVA)	$S^{*F}_{ij,t}$	Maximum capacity need upgrading of Feeder (MVA)
C^{PV}_i	New investment cost for PV i ($/M)	ΔS^F_{min}	Capacity ramp-up limit for Feeder (MVA)
$F^*_{ij,t}$	Standard section of Feeder in planning year t (mm^2)	$S^F_{max,ij,t}$	Maximum capacity limit of standard Feeder (MVA)
J	Current density at thermal limit (A/mm^2)	$S^{*S}_{i,t}$	Maximum capacity need upgrading of Substation (MVA)
M	Big number used maximum limit of variables in MIP and MINLP	ΔS^S_{min}	Capacity ramp-up limit for Substation (MVA)
U_{max}	Maximum voltage limit at bus (pu)	$S^S_{max,i,t}$	Maximum capacity limit of standard Substation in planning year t (MVA)
U_{min}	Minimum voltage limit at bus (pu)	k_p	Variation factor of the price of electricity
ΔP^{PV}	Active power ramp-up limit for PV (MW)	D_S	Total day per season
$k^{PV}_{s,h}$	Output power factor of PV	ξ	Emission coefficient of traditional energies
		β	Emission tax

The costs arising from emission taxes can be reduced due to lower emissions when PV using renewable energy resources is substituted for the traditional energies in DS (TCO_t). This part is negative and shown in Eq. (7) with ξ is emission coefficient and β is emission tax that may be enforced by the government.

$$TCO_t = \sum_{i=1}^{N_{PV}} \sum_{S=1}^{S_S} \sum_{h=1}^{H} \beta.\xi.D_S.P^{PV}_{i,t,s,h} \qquad \forall i \in N_{PV}, t \in T, s \in S_S, h \in H \tag{7}$$

The residual value of equipments at the end of the planning period is presented in Eq. (8) and it is usually evaluated basic on the current market conditions. Hence, the residual value is the present value and it is calculated at base year in objective function.

$$RN_t = \frac{(t^F_{kh} - T_F)}{T_F}.CF_t + \frac{(t^S_{kh} - T_S)}{T_S}.CS_t + \sum_{i=1}^{N_{DG}} \frac{(t^{PV}_{kh} - T_{PV})}{T_{PV}} C^{PV}_i.P^{PV}_{i,t} \qquad \forall i \in N_{PV}, t \in T \tag{8}$$

2.1.2 The Constraints

**) Constraints for Power Flow*

The output power of PV fluctuates by solar radiation so it is also determined by time of the day and season in year. Hence, an AC nonlinear power flow model is used in this stage as represented in (9). With this constraint, the influences of reactive power to calculation power and voltage losses in DS are considered so results of proposed model are more accurate.

$$P^{S}_{i,s,t,h} - PD_{i,s,t,h} + P^{PV}_{i,s,t,h} = \sum_{j=1}^{N} |Y_{ij,t}|.|U_{i,s,t,h}|.|U_{j,s,t,h}|.\cos(\theta_{ij,t} - \delta_{j,s,t,h} - \delta_{i,s,t,h})$$

$$Q^{S}_{i,s,t,h} - QD_{i,s,t,h} = -\sum_{j=1}^{N} |Y_{ij,t}|.|U_{i,s,t,h}|.|U_{j,s,t,h}|.\sin(\theta_{ij,t} - \delta_{j,s,t,h} - \delta_{i,s,t,h})$$

$$\forall i, j \in N, s \in SS, h \in H, t \in T \tag{9}$$

where, $P^{PV}_{i,t,s,h}$ is the output power of PV that changes depending on the primary energy introduced in (10).

$$P^{PV}_{i,s,t,h} = P^{PV}_{i,t}.k^{PV}_{s,h} \tag{10}$$

**) Capacity Limited Constraints of PV*

These constraints allow the selection of PV capacity in its limits at each node, and it ensures annually upgrading power corresponding to the equipment parameters as shown in (11).

$$0 \le P^{PV}_{i,t} \le P^{PV}_{i,\max}; P^{PV}_{i,t} = P^{PV}_{i,t-1} + \Delta P \qquad \forall t \ge 1, i \in N_{PV}, t \in T \tag{11}$$

**) Upgrading Section Constraints of Feeder*

This constraint is imposed to limit the loading of feeders and these limits are taken into the consideration of new feeder investments. Thus, the feeder upgrading constraints and upgrading power satisfying equipment parameters are shown in (12). A step increase of feeder capacity at year t ($\Delta.S^{F}_{ij,t}$) is set when the capacity value is equal or greater than the capacity limit used at year t−1.

$$S^{max}_{ij,t} \le (S^{*F}_{ij,t-1} + \Delta S^{F}_{\mathrm{ij},t}); \Delta S^{F}_{\mathrm{ij},t} \ge \Delta S^{F}_{\min}.\alpha_{\mathrm{ij},t}; \Delta S^{F}_{\mathrm{ij},t} \le M.\alpha_{\mathrm{ij},t} \quad \forall t \ge 1, ij \in N, t \in T \tag{12}$$

Then, the feeder capacity needs to meet in order to supply power to the load and the upgrading section is selected present in (13) with current density J.

$$S^{*F}_{ij,t} = S^{*F}_{ij,t-1} + \Delta S^{F}_{\mathrm{ij},t}; \quad F_{\mathrm{ij},t} \ge \frac{S^{*F}_{ij,t}}{\sqrt{3}U_{dm}.J}.\alpha_{ij,t} \quad \forall t \ge 1, ij \in N, t \in T \tag{13}$$

**) Addition Capacity Constraints for Substation*

These constraints allow to maximize the use of existing substations capacity and to satisfy upgrading power corresponding to the equipment parameters. A substation capacity addition step size ($\Delta.S^{S}_{ij,t}$) is used as in Eq. (14) with the maximum and minimum allowable capacity which can be upgraded.

$$S^{max}_{i,t} \le (S^{*F}_{i,t-1} + \Delta S^{S}_{\mathrm{i},t}); \Delta S^{S}_{\mathrm{i},t} \ge \Delta S^{S}_{\min}.\gamma_{\mathrm{i,t}}; \Delta S^{S}_{\mathrm{i},t} \le M.\gamma_{\mathrm{i,t}} \quad \forall t \ge 1, i \in NS, t \in T \tag{14}$$

*) *Constraints of Nodal Voltage Limited*
The technical requirement constraints of limited nodal voltage are given in Eq. (15). The voltages at substation nodes are assumed constantly.

$$\begin{aligned} U_{\min} \le |U_{i,s,t,h}| \le U_{\max} \quad & \forall i \in N_L, s \in SS, t \in T, h \in H \\ |U_{i,s,t,h}| = constant \quad & \forall i \in N_S, s \in SS, t \in T, h \in H \end{aligned} \tag{15}$$

The decision variables of models include real and binary variables so calculation results must be corrected by the standard equipment in reality and used as parameters in the second stage.

2.2 The Mathematical Model of Second Stage

This stage takes the input parameters obtained from the first stage as the additional capacity of substations, upgrading section of feeders, installation location and period of PV. Then, it determines the PV capacity within predefined bounds.

2.2.1 Objective Function
The model has objective function similar to the first stage (J_1) with upgrading variables of feeders ($F_{ij,t}$) and substations ($\Delta.S^S_{ij,t}$) are replaced by the equipment parameters obtained from the first stage. Hence, the equations of objective function are presented as (16) and decision variable PV power is ($P^{PV}_{i,t}$).

$$CF_t = \sum_{i=1}^{N} \sum_{j=i+1}^{N} L_{ij}(C^{FF}.\alpha_{\mathrm{ij},t} + C^{FC}.F^*_{ij,t}); CS = \sum_{i=1}^{NS}(C^{SF}.\gamma_{i,t} + C^{SC}.S^{*S}_{i,t}) \tag{16}$$

2.2.2 The Constraints
*) *Constraints for Power Flow and Nodal Voltage Limited*
These constraints are similar to the first stage and presented on Eqs. (9) and (15).
*) *Capacity Limited Constraints of Feeder and Substation*
To ensure the upgraded feeders are not overloaded by thermal limits, the load flow on feeder need satisfy and substation capacity must satisfy as Eq. (17).

$$\begin{aligned} S^F_{ij,t,s,h} \le S^F_{\max.ij,t} \quad & \forall t \ge 1, ij \in N, t \in T, s \in SS, h \in H \\ S^S_{i,t,s,h} \le S^S_{max.i,t} \quad & \forall t \ge 1, i \in NS, t \in T, s \in SS, h \in H \end{aligned} \tag{17}$$

*) *Capacity Limited Constraints of PV*
The investment location and period of PV were determined from the first stage so these constraints allow the selected PV capacity according to new limits as (18).

$$0 \le P^{PV}_{i,t} \le P^{*PV}_{\max.i}; P^{PV}_{i,t} = P^{PV}_{i,t-1} + \Delta P \qquad \forall t \ge 1, i \in N_{PV}, t \in T \tag{18}$$

The proposed comprehensive plan includes a MINLP model in first stage and NLP model in second stage. The calculation program is made in GAMS environment used MINOS solver [10] to find out an optimal solution.

3 Results and Discussions

3.1 Diagram and Parameters of Distribution Systems

A 7-bus and 22 kV voltage radial diagram is investigated in this research as Fig. 1 and is connected to 110 kV transformer substation. The total active power and reactive power at the base year are 8838.0 kW and 7309.2 kVAR, respectively.

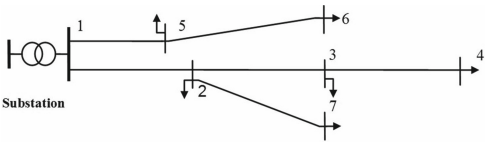


Fig. 1. Diagram of test distribution system

3.2 Assumptions in Analysis

This research utilizes some economic and technical assumptions:

+ The planning period is 5 years and the annual developing rate of load demand is constant, 10 % per year. At all of the load locations the typical characteristics of load demands for four seasons+ The constructing cost of 110 kV substation including fixed costs and variable costs is 0.2 M$ and 0.05 M$/MVA, respectively [8]. Similarly, the upgrading costs of 22 kV feeders consist of 0.15 M$/km and 0.001 M$/MVA.km. The assumption life of feeder is 20 year.

+ The PV sources are used in this research with the corresponding capital costs to be 3.0 M$/MW. The average operation and management (O&M) costs is 5 $/kWh and the life of PV is 30 years.

+ The energy prices purchasing from EM are real-time prices.

+ The PV is manufactured in compact modules occupying small spaces and time to install is short. Hence, the installing areas are not limited and PV can be selected to install at all of load locations. Similarly, areas for upgrading of substation transformers and feeders are not limited.

+ The constraints of load nodes voltage limited allow change from 0.9 pu to 1.1 pu, and it should be 1.05 pu at substation node.

+ The decided variables in the model are continuous in order to reduce the complexity of the model. Hence, they should be rounded to match real equipments.

3.3 The Output Power Characteristics of PV

The power output of PV depends on the intensity of solar radiation and its performance. The power factor of PV calculated basing on the given solar radiation intensity is presented as Fig. 2. In contrast, gas turbine does not depend on the nature of the primary energy source so its output power can be constant.

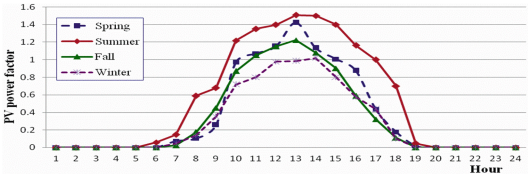


Fig. 2. The output power characteristics of PV

3.4 Analysis Results and Discussions

The feasibility of the proposed model and the efficiency of PV are investigated in two cases. Case A is calculated when PV is not considered. While, case B integrates PV in the researching model to plan the DS.

The calculation results showed that both cases are upgraded the substation with 10 MVA capacity. However, case A is just upgraded at the first year and the investment to upgrade substation in case B is deferred to 3rd year because of the load demand increasing in the future is provided by PV. Similarly, in the case A, 3 feeders need be upgraded in the time from 2nd year to 5th year as represented in Table 2. The feeder 2–3 in case B is not upgraded during the planning period. Only feeders 1–2 and 1–5 are invested to upgrade at 3rd year and 5th year.

Table 2. Feeders upgrading decisions

Feeder	Feeder section upgrading in eyear t(mm^2)									
	1	*2*	*3*	*4*	*5*	*1*	*2*	*3*	*4*	*5*
	Case A					*Case B*				
1–2	-	70	-	-	-	-	-	70	-	-
2–3	-	-	-	-	70	-	-	-	-	-
3–4	-	-	-	-	-	-	-	-	-	-
1–5	-	-	-	50	-	-	-	-	-	50
5–6	-	-	-	-	-	-	-	-	-	-
2–7	-	-	-	-	-	-	-	-	-	-

Table 3 presents the optimal investment decisions of proposed planning model for PV. The total of investment capacity during planning time is 4.0 MW equivalent to 45.2 percent of load demands at base year. The PV are invested and selected that location of these sources are far from substation. Therefore, high economic and technical efficiencies are gained. Economic indicators are compared between case B and case A as shown in Table 4. The case B holds a better economic indicator. Costs for the investment of PV and feeders, substations upgrading are more expensive than those in case A about 9.27 M$ due to a very high cost of PV investment but the life of PV is 30 years. However, O&M and electrical energy

expenses that calculated at the base year have been decreased about 0.47 M$ because of very low O&M expenses of PV. Therefore, total life cycle cost of case B is lower than these of case A by 0.13 M$, equal to 0.88 %. The technical indicators of DS are also improved when PV is integrated on DS planning. The electrical energy loss always reduces during planning period as represented in Fig. 3. At just 1st planning years the electrical energy loss is reduced 0.74 %, the corresponding 322.92 MWh. This value decreases in 4th planning years and it is only 0.25 % because the feeders upgraded of case A decrease the resistors of DS. The total of electric energy purchased from markets is also decreased 70,700.0 MWh corresponding to 18,382.0 tons CO2 emission of traditional sources [8], which contributes to the decrease of environmental pollutions. The capacity on the feeders in case B is lower than in case A due to supported by PV that installed at load nodes. Therefore, the voltage loss of the system reduces and voltage profiles at the all bus are improved during the calculation time. In particular, the load node having the biggest support is 4-bus as shown on Fig. 4. This bus voltage profile increased 1.3 % at 18th hour in 1st planning year. The 2nd year selected to upgrade feeder 1–2 so the difference of voltage profiles in 2 cases are dropped and then they increase at next year depending on the rise of load demand.

Table 3. PV investment decided

Bus	PV capacity invested in year t (MW)					Bus	PV capacity invested in year t (MW)				
	1	*2*	*3*	*4*	*5*		*1*	*2*	*3*	*4*	*5*
2	-	-	-	-	-	5	-	-	-	-	-
3	-	1.0	-	-	-	6	-	-	1.0	-	-
4	1.0	-	-	-	-	7	-	-	-	1.0	-

Table 4. Economic indicators comparison

No	Cost	Case A	Case B	Comparison B and A	Note
1	Total life cycle cost (M$)	14.84	14.71	-0.13	Total life cycle costs is reduced - 0.88%
2	Feeder and Substation upgrading cost (M$)	3.52	2.61	-0.91	
3	O&M and Electrical energy cost (M$)	14.32	13.85	-0.47	
4	Investment PV cost (M$)	0	10.18	10.18	

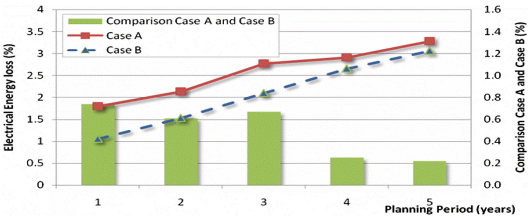


Fig. 3. Comparison electrical energy loss between Case B and Case A

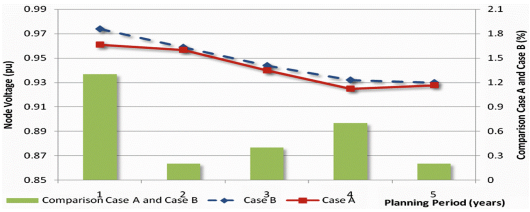


Fig. 4. Comparison voltage of 4-bus between Case B and Case A

4 Conclusions

Recently, the DS planning has been changed significantly by the impacts of PV and environmental policies. The effect of DS is improved by PV as the enhancement of flexibility and reliability, bus voltage improvement, reduction of transmission cost and power loss as well as the reduction of environmental pollution. However, the investment cost of PV is usually expensive and power of these sources that used renewable energy has natural variability according to the primary energy so the planning and operation calculation of DS will be more difficult. Therefore, this study proposed a new two-stage optimized model that is integrated PV in DS planning problem. In this model, the equipment sizing and time-frame required for upgrading equipment of DS well as selection technologies and power variable constraints of PV can be determined. The objective function is minimizing total life cycle cost of the investment project. The calculated results showed that the proposed model is suitable in DS planning calculation and the planning together with using PV provided better economic and technical indicators. The total life cycle cost of planning project that integrated PV always reduces. The effects of PV on DS are either on loss reduction or feeders and substations capacity deferment. The power and electrical energy losses also decrease, and furthermore the voltage profiles of nodes are always improved.

References

1. Albadi, M.H., El-Saadany, E.F.: The role of distributed generation in restructured power systems. In: Proceedings of NAPS 2008, Canada (2008)
2. Muneer, W., Bhattacharya, K., Cañizares, C.: Large-scale solar PV investment models, tools and analysis: the ontario case. IEEE Trans. Power Syst. **26**, 2547–2555 (2011)
3. Medina, A., Hernandez, J.C., Jurado, F.: Optimal placement and sizing procedure for PV systems on radial distribution systems. In: International Conference on Power System Technology, China (2006)
4. Ackermann, T., Andersson, G., Söder, L.: Distributed generation: a definition. Electr. Power Syst. Res. **57**, 195–204 (2001)
5. Wong, S., Bhattacharya, K., Fuller, J.D.: Comprehensive framework for long-term distribution system planning. In: Proceedings of IEEE PES Annual General Meeting USA (2007)

6. Algarni, A.A.S., Bhattacharya, K.: Novel approach to disco planning in electricity markets: mathematical model. In: Proceedings of PSCE 2009. IEEE/PES (2009)
7. El-Khattam, W., Hegazy, Y., Salama, M.: An integrated distributed generation optimization model for distribution system planning. In: Power Engineering Society General Meeting. IEEE (2005)
8. Wong, S., Bhattacharya, K., Fuller, J.D.: Electric power distribution system design and planning in a deregulated environment. IET Gener. Trans. Distr. **3**(12), 1061 (2009)
9. Thang, V.V., Thong, D.Q., Khanh, B.Q.: A new model applied to the planning of distribution systems for competitive electricity markets. In: Proceedings of DRPT 2011, China (2011)
10. Rosenthal, R.E.: GAMS - A User's Guide. GAMS Development Corporation, Washington (2010)

The Prediction of Succinylation Site in Protein by Analyzing Amino Acid Composition

Van-Minh Bui[1(✉)] and Van-Nui Nguyen[2]

[1] Danang College of Technology, The University of Danang, Danang, Vietnam
vanminha9@gmail.com, bvminh@ud.edu.vn
[2] University of Information and Communication Technology, Thai Nguyen, Vietnam
nvnui@ictu.edu.vn

Abstract. Protein Succinylation is a kind of post-translational modification (PTM) where a succinyl group is attacked to a lysine residue of a protein molecule. Recent findings have demonstrated the important role of Succinylation in not only taking part in various biological processes but also associating with many diseases. There are many practical methods to identify succinylation sites but an expensive cost and time-wasting should be considered. The lack of research in structure and characteristic of protein will limit to understand and discover significantly. Therefore, this work aims to focus on develop a bioinformatics method for investigating Succinylation site based on the amino acid composition and physicochemical properties. Various features were investigated in this study, including 20 Binary coding, amino acid composition (AAC), amino acid pair composition (AAPC), solvent-accessible surface area (ASA), amino acid substitution matrix (Blosum62), and position-specific scoring matrix (PSSM). Evaluation by five-fold cross validation indicated that the selected features were effective in the identification of Succinylation sites. The model constructed from hybrid features, including BLOSUM62 and PSSM, yielded the best performance with sensitivity, specificity, accuracy and MCC measurements of 0.66, 0.68, 0.67 and 0.32, respectively.

Keywords: Succinylation · Post-translational modification · Support vector machine

1 Introduction

Post-translational modification (PTM) hold a vital role in forming proteomic diversity. Succinylation is a novel PTM which occur in both prokaryotic and eukaryotic cells. This PTM was constructed by adding succinyl group (–CO–CH2–CH2–Co–) to lysine residues of substrate proteins. Compare to other lysine PTMs (acetylation, methylation), succinylation affect directly to contribute to regulate protein structure, function and physicochemical properties. The first found in Succinylation happened at the active site of homoserine trans-succinylase [1]. With mass spectrometry and an antibody-based affinity enrichment strategy and MS-based proteomics, number of succinylation sites was identified in bacteria in 2011 by Zhang [2]. Moreover, later work gradually detected

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_67

recognized plenty of succinylation sites in both animal tissues [3] and on histones [4]. Until now, the number of identified succinylation protein was limited by the drawback of practical methods. As identification of succinylated substrate proteins can explore useful wisdom to understand the molecular mechanism of succinylation in biological systems and overcoming a major hurdle in this field. To tackle the big problems in processing a large-scale data, proteomic technologies with labor-intensive and time-consuming should be considered and replaced by the computation approaches. With the developing of dominant bioinformatics technologies, these approaches are utilized for prediction of individual lysine residues to succinylation. In this regard, the computational approaches were effectively and accurately adopted to identify the succinylated sites by analyzing amino acid compositions. Based on the In silico characterization of protein, some sequential and structural features including amino acid composition (AAC), amino acid pair composition (AAPC), position specific scoring matrix (PSSM), BLOSUM62 and accessible surface area (ASA) were applied to discriminate between the Succinylation sites and non-succinylation sites.

2 Materials and Methods

2.1 Data Preprocessing of Training Set and Independent Test Set

Training data was collected from 4 resources (Table 1). A majority of the experimental data used in this study was obtained from the NNF Center for protein Research, which stores information on experimentally verified Succinylatied lysines in 4 species based on antibody-based affinity enrichment of succinylated peptides followed by strong cation exchange (SCX) chromatography [5]. There are 2424 succinylated proteins composed of 6542 succinylated sites. An additional set of data was collected from CPLM [6], and other published literature. From the CPLM database, we extracted 2439 Succinylated sites (positive dataset) in 850 proteins, while the rest of the lysine residues in these proteins were considered as non-succinylated lysines (negative dataset). Using high-accuracy nano liquid chromatography-tandem mass spectrometry combined with affinity purification [7], we identified 1931 lysine succinylated peptides matched on 642 proteins. Finally, our training data also was gathered from other published study (PMID: 26663479) with 636 succinylated site from 281 proteins. In order to generate the training dataset, the window length of $2n + 1$ was used to extract sequence fragments centering at the Succinylated or non-Succinylated lysine residues and containing n upstream, as well as n downstream, flanking amino acids. After extracting the sequence fragments with $2n + 1$ window length ($n = 10$) from 4 resources, the training data included 10948 positive data and 81428 negative data. Similar to training data, the testing data also was collected from 4 public practical studies as the Table 2 (PMID: 26828863, 24315375, 25377623 and 25363132). Then, testing data was separated into small fragments with same window size in training data and they were finally 3351 succinylated sites (positive testing data) and 18967 non-succinylated sites (negative testing data) from 1187 proteins.

Table 1. Data statistics of experimentally verified Succinylation in training dataset

Data Resources	Number of proteins	Succinylation sites (Positive data)	Non-Succinylation sites (Negative data)
Brian T. Weinert et al., 2013 (PMID: 23954790)	2424	6542	-
CPLM database (PMID: 24214993)	850	2439	-
Jianyi Pan et al., 2015 (PMID: 26369940)	642	1931	-
Yuta Mizuno et al., 2015 (PMID: 26663479)	281	636	-
Combined dataset (non-redundant)	**3839**	**10948**	**81428**

Table 2. Data statistics of Succinylation in independent dataset

Data Resources	Number of proteins	Succinylation sites (Positive data)	Non-Succinylation sites (Negative data)
Weibo Jin et al., 2016 (PMID: 26828863) [19]	200	339	-
Matthew J. Rardin et al., 2013 (PMID: 24315375) [20]	221	992	-
Xiaolong Li et al., 2014 (PMID: 25377623)[21]	147	423	-
Longxiang Xie et al., 2015 (PMID: 25363132)[22]	619	1597	-
Combined dataset (non-redundant)	**1187**	**3351**	**18967**

The purpose of our study concentrated on the sequence-based analysis of substrate site specificity of Succinylated lysines. For detecting the training model, the five-fold cross validation is adopted to evaluate the probability of distinguishing between the Succinylated and non-Succinylated lysine residues. After evaluating training set by using the cross-validation, the best model was further chosen by using the independent test set. Nevertheless, there is the probability of existence of the homologous sequences in both the training and independent testing set. As for classification, the over fitting of some sequences is in both a training and independent set, it may lead to overestimated prediction performance of the trained model. Therefore, removing the homologous sequences within both the training and independent test sets is required to ensure the high reliabilities. CD-HIT package [8] is a valuable tool for clustering proteins that has a similarity threshold and one instance sequence was chosen to represent for each cluster. Being founded on the CD-HIT advantages, homologous sequences were efficiently removed with different thresholds from 40 % to 100 %. As shown in Table 3, there were appropriate 1478 proteins to be removed on training data corresponding threshold 40 % equal to 2420 proteins. For testing data, we did not utilize CD-HIT to remove similarity sequences. Moreover, to avoid the redundant data in cases of finding multiple records of the same site position and accession number, our database also was solved this problem. The number of distribution between succinylation and non-succinylation data is seriously unbalanced, we randomly selected twice negative dataset (11158 in 57770) to combine with positive dataset (5579) in training data. For independent testing, we still keep all

Table 3. Data statistics after using CD-HIT

Threshold	Number of proteins	Succinylation sites (Positive data)	Non-Succinylation sites (Negative data)
Full data	3839	10218	81428
90 %	3475	9113	78506
80 %	3217	8155	74222
70 %	2965	7292	68842
60 %	2793	6771	65309
50 %	2627	6171	61731
40 %	2420	5579	57940

the positive and negative to evaluate the training model. Finally, the numbers of positive and negative fragment sequences were 5579 and 11158 for the training data set and 3051 and 16304 for the independent test set, proportionally.

2.2 Features Investigation

To build the predictive models, support vector machine (SVM) was adopted for defining S-succinylation and non-Succinylation sites on proteins based on several selected features such as 20D binary code (AA), AAC, BLOSUM62, ASA, AAPC, PSSM and PWM. The fragment sequences, which contained the Succinylation cysteine in the middle, were extracted with window size equal to 21 from positive and negative data [9]. An orthogonal binary coding mechanism is one of the most popular coding methods to convert amino acids into numeric vectors, called 20D binary code. For example, Alanine (A) was encoded as "10000000000000000000", Cysteine (C) was encoded as "01000000000000000000" and so on. The number of feature vectors was $(2n + 1) \times 20$ to represent the flanking amino acids surrounding the Succinylation sites. There are a total of k vectors $\{x_i, i = 1, 2, \ldots, k\}$ corresponding with number of k fragment sequences in the training data. To classify the positive and negative cysteine, label was applied for each vector to mark the class of its corresponding protein. For composition of amino acid around the Succinylation sites, the vector x_i had 21 elements for AAC and 441 elements for AAPC. Excepting the occurrence frequencies of 20 amino acids in a sequence fragment, there is a probability of existence of either rarely special amino acid or non-existing "-" residues which was used to represent for less than 21-mer fragment sequences at N- or C-terminus. The BLOSUM62 [10] was built based on the alignments of amino acid sequences that the no more than 62 % identity between two peptide sequences with 21 amino acids. PSSM profiles were generated from PSI-BLAST [11] against non-redundant sequences of Succinylation sites and this matrix of score values can represent the multiple sequence alignment of proteins which may have similar structures with different amino acid compositions. Extracting from the PSSM profile, the matrix of $(2n + 1) \times 20$ elements had rows centered on substrate site, where $2n + 1$ represented the window size and 20 represented the position specific scores for each type of amino acid. The structural feature

of ASA has been investigated based on the accessible ability of a side-chain of amino acid on the surface of a protein that experience post-translational modification [12]. RVP-Net [13,14], an effective tool, was applied to compute the ASA value from the protein sequence due to lacking of most of experimental Succinylated protein tertiary structures in PDB [15]. Based on information about their neighborhood, RVP-Net could predict the real ASA of residue by using a neutral network. The possible mean absolute error which was defined as the absolute difference between the predicted and experimental values of relative ASA per residue [14], was $18.0 - 19.5\,\%$, for each measurement. The value of ASA was the percentage of the solvent-accessible area of each amino acid on the protein. The input data of RVP-Net were the full-length protein sequences to compute the ASA value of all of the residues. Then, the ASA values of amino acids around the Succinylation sites were extracted and normalized to be from zero to one. The hybrid features were generated by combining from two or more single features. Based on the performing of each feature, the single features were selected to gain the highest predictive accuracy. Before using classification, the data need to be scaled in range $[-1, 1]$ to enhance effectively the results.

2.3 Model Learning and Evaluation

The positive and negative of training data set are built the predictive model by using on SVM. Based on binary classification, the way using a kernel function is to transform into a higher dimensional space from the input samples and then find a hyper-plane to discriminate the two classes with maximal margin and minimal error. In our study, a public SVM library (LIBSVM) [16] was implemented the predictive model for distinguish the Succinylated sites from non-Succinylated sites. The radial basis function (RBF) was adopted as the kernel function for learning SVM classifier. There are two supporting factors to enhance the performance with gamma and cost. The RBF kernel is determined by gamma parameter while the cost parameter controls the hyper-plane softness. Each feature was predicted by LIBSVM library, followed by generating the best feature as an input vector for the second-layered SVM based on the values of previous estimated probability from each SVM classifier. $K(S_i, S_j) = \exp(-\gamma||S_i - S_j||^2)$ (1) To choose the best final model, the five-fold cross validation was carried out for each different feature to evaluate the predictive performances. The training data was divided into five approximately equal sized subgroups. The ratio of test set and training set are 1:4 and the cross-validation process was repeated five times. The five validation results were then combined to generate a single estimation. Obviously, one of the benefits of cross-validation evaluation is to improve the reliability of evaluation because all original data, including training and testing data set, were considered in general and each subset was tested only once [17]. To gauge the effectively predictive performance of training model, the following measures were used such as sensitivity (Sn), specificity (Sp), accuracy (Acc) and Matthews Correlation Coefficient (MCC):$\mathrm{Sn} = \mathrm{TP}/(\mathrm{TP} + \mathrm{FN})$ (2) $\mathrm{Sp} = \mathrm{TN}/(\mathrm{TN} + \mathrm{FP})$ (3) $\mathrm{Acc} = (\mathrm{TP} + \mathrm{TN})/(\mathrm{TP} + \mathrm{FP} + \mathrm{TN} + \mathrm{FN})$ (4) $\mathrm{MCC} = ((\mathrm{TP} * \mathrm{TN}) - (\mathrm{FN} * \mathrm{FP}))/((\mathrm{TP} + \mathrm{FN})(\mathrm{TN} + \mathrm{FP})(\mathrm{TP} + \mathrm{FP})(\mathrm{TN} + \mathrm{FN}))$

(5) where TP, TN, FP and FN represented the numbers of true positives, true negatives, false positives and false negatives, respectively. Finally, after selecting the best predictive model with the highest accuracy, an independent testing was carried out to test the final model with best performance in cross-validation evaluation. Four mentioned measurement units was evaluated the predictive capability of our method. Sensitivity is the percentage of correct predictions from positive case (Succinylated residues) while specificity represents for that of negative case (non-Succinylated lysines). Accuracy reflects the overall proportion of correctly predicted Succinylation and non-Succinylation peptides. The MCC is considered as a more reliable measure of the quality of binary classification. The range of MCC is from −1 to +1. A coefficient of positive value represents a perfect prediction while the value −1 and 0 are revealed for opposite and random predictions, respectively.

3 Results and Discussion

3.1 Influence of Flanking Amino Acid Composition Around Succinylation Sites

This study concentrates on analyzing frequent occurrences of 20 amino acid surrounding the lysine resides to identify the Succinylated sites. Through sequence-based investigation, the frequency of twenty amino acids surrounding the Succinylation sites revealed the amino acid composition that forms the substrate environment for protein Succinylation. The distribution of twenty amino acids between Succinylation and non-Succinylation site has few interestingly different things as shown in Fig. 1. The Fig. 1A represent, at Succinylation sites, A (Alanine), R (Arginine), and D (Aspartic acid) residues occur at a higher frequency. Combining with Fig. 1B, the frequency of each amino acid was described clearly at each position (window size = 21). At the position 0, the number of amino acid Lysine made up the highest result and the dispersion of two amino acids R – Arginine and D – Aspartic acid got attention in its frequency. Interestingly, based on the proportion of 20 amino acid's occupancies with over 5 % near Succinylation sites (Fig. 2A) and the distribution of 20 amino acids at each position (Fig. 2B), noticeable amino acids are E with over 7 % (at position −4, −3, 1, 3 and 4), G with around 6.5 % (at position −1, 1, and 2), L with near 9 % (at position −2, and −1), and R with over 5 % (at position −8, −7, 1, 7 and 8). TwoSampleLogo [18] is an effective web-based tool to detect statistically noteworthy differences in position-specific symbol compositions between Succinylated and non-Succinylated data set. The center position of fragment sequences is lysine residue and range value is from −10 to 10 corresponding to the position of flanking amino acid. Comparing between 5579 Succinylation sites and 57770 non-Succnylation sites, the positively charged amino acids including Lysine (K) and Arginine (R) made up the highest radio at positions −10 −6, 1, and 4 10 (with p-value ¡ 0.01). Obviously, the rest of positions where is around Succinylation sites are lack of positively charged residues at positions −5 −1, 2,



Fig. 1. Influence of twenty amino acid around Succinylation sites

and 3 and be replaced by the appearance of polar group (Tyrosine – Y, Threonine – T, and Glycine – G) and non-polar group (Leucine – L, Phenylalanine – F, Valine – V, Alanine – A, Isoleucine – I, Methionine – M, and Tryptophan –W) at positions −5 −1. In addition, the existence of negatively charged residues (Glutamic acid – E and Aspartic aid – D) was at positions −2, 1, and 2. By contrast with positive dataset, the beingness of positively charged residue in negative one was at position −3 2 near the non-Succinylation sites. The analysis indicated that the distance among amino acid characteristic in sequence play a vital role in discriminate between Succinylation sites and non-Succinylation ones. To estimate carefully for the composition of amino acids in sequences, RVP-Net algorithm was adopted to evaluation the correlation between Succinylation sites and solvent accessible surface area (ASA). Computed by the RVP-Net tool, ASA was also adopted as an attribute for the identification of Succinylation sites. To discover how amino acids flanking the Succinylation and non-Succinylation sites may differ in their interaction with solvents, a comparison was performed using the average proportion of ASA in the 21-mer window (−10 +10) as illustrated in Fig. 2. Amino acids surrounding the Succinylation sites exhibit lower ASA compared to those around non-Succinylation sites. A strong evidence for hydrophilicity at the Succinylated substrate sites was found because the average percentage of ASA values of the flanking residues was lower than non-Succiylation lysine; especially at positions −5 −1. Thus, hydrophilic amino acids flanking lysine residues could be a factor to determine their modification.

3.2 Cross-Validation of Characteristics for Flanking Amino Acids and Succinylation Sites

The four measurement units (includig Sn, Sp, Acc, MCC) were applied to evaluate the performance of predictive models. As shown in Table 4, the SVM-based model with ASA had the lowest predictive accuracy at 0.58 and relatively lower sensitivity, specificity, and MCC at 0.57, 0.59 and 0.15, respectively.

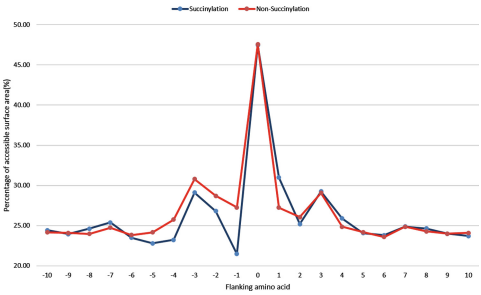


Fig. 2. Comparison of solvent-accessible surface area between Succinylation and non-Succinylation sites

By contrast, the analysis indicated that BLOSUM62 feature was regarded as the best one for training a model with sensitivity at 0.62, specificity at 0.67, accuracy at 0.66 and MCC at 0.28. Furthermore, PSSM and 20D Binary code were two potential features to combine with the best feature BLOSUM62 to enhance the performance. Additionally, predictive mode based on the hybrid combination of BLOSUM62 with PSSM provided the best predictive performance in Sn, Sp, Acc and MCC at 0.66, 0.68, 0.67 and 0.32, respectively. Obviously, the rate of sensitivity of hybrid feature was enhanced a little better than that of BLOSUM62. Therefore, the hybrid feature of PSSM and BLOSUM62 were selected as the training feature for the construction of a potential SVM model.

Table 4. Five-fold cross validation results on the SVM model trained with various features

Training features	**Sn** proteins	**Sp**(Positive data)	**Acc**(Negative data)	**MCC**
20D Binary code (AA)	0.64	0.64	0.64	0.27
BLOSUM62	0.62	0.67	0.66	0.28
Amino acid composition	0.59	0.61	0.60	0.19
Amino Acid Pair Composition (AAPC)	0.58	0.64	0.62	0.21
Accessible Surface Area (ASA)	0.57	0.59	0.58	0.15
Position-specific scoring matrix (PSSM)	0.61	0.67	0.65	0.26
BLOSUM62 + AA	0.63	0.67	0.66	0.29
BLOSUM62 + PSSM	**0.66**	**0.68**	**0.67**	**0.32**

3.3 Evaluation of Succinylation's Predictive Models Using Independent Testing Set

An independent test set of succinylation was adopted to evaluate the effectiveness of hybrid SVM model between BLOSUM62 and PSSM that achieved the best accuracy in five-fold cross-validation. In this investigation, we used 3051 positive data and 16304 negative data as independent test data set. As summarized in Table 4, there are widely differences in the ratio of sensitivity, specificity,

Table 5. The detailed independent testing results between our methods

Features	TP	FP	TN	FN	Sn	Sp	Acc	MCC
B62	1758	6760	9544	1293	0.58	0.59	0.58	0.12
PSSM	1758	6548	9756	1293	0.58	0.6	0.59	0.13
B62 + PSSM	1765	6322	9983	1286	0.58	0.61	0.61	0.14

accuracy and MCC between three models at 0.58, 0.59, 0.58 and 0.12 respectively for BLOSUM62 model and 0.58, 0.6, 0.59, 0.13 respectively for PSSM model and, last one, 0.58, 0.61, 0.61, 0.14 respectively for hybrid model between BLOSUM62 and. In general, the evidence shows that hybrid SVM models (BLOSUM62 + PSSM) resulted in better predictive performance than the single SVM model (Table 5).

4 Conclusion

In this study, we focus on analyzing amino acid compostion to identified the Succinylation sites. Through TwoSampleLogo, the analysis of position-specific amino acids composition between Succinylation and non-Succinylation site figured out the relationship between flanking amino acids with Succinylation lysine residue. Moreover, this investigation also mentioned the ability of solvent-accessible surface of amino acid surrounding Succinylation sites have a tendency to lower than that around non-Succinylation sites. According to the results of five-fold cross-validation, the hybrid feature of BLOSUM62 and PSSM is estimated as the best feature with the highest proportion of sensitivity, specificity, accuracy and MCC.

References

1. Rosen, R., et al.: Probing the active site of homoserine trans-succinylase. FEBS Lett. **577**(3), 386–392 (2004)
2. Zhang, Z., et al.: Identification of lysine succinylation as a new post-translational modification. Nat. Chem. Biol. **7**(1), 58–63 (2011)
3. Du, J., et al.: Sirt5 is a NAD-dependent protein lysine demalonylase and desuccinylase. Science **334**(6057), 806–809 (2011). (New York, N.Y.)
4. Xie, Z., et al.: Lysine succinylation and lysine malonylation in histones. Mol. Cell. Proteomics (MCP) **11**(5), 100–107 (2012)
5. Weinert, B.T., et al.: Lysine succinylation is a frequently occurring modification in prokaryotes and eukaryotes and extensively overlaps with acetylation. Cell Rep. **4**(4), 842–851 (2013)
6. Liu, Z., et al.: CPLM: a database of protein lysine modifications. Nucleic Acids Res. **42**(Database issue), D531–D536 (2014)
7. Pan, J., et al.: Global analysis of protein lysine succinylation profiles and their overlap with lysine acetylation in the marine bacterium vibrio parahemolyticus. J. Proteome Res. **14**(10), 4309–4318 (2015)

8. Li, W., Godzik, A.: Cd-hit: a fast program for clustering and comparing large sets of protein or nucleotide sequences. Bioinformatics **22**(13), 1658–1659 (2006)
9. Lee, T.Y., et al.: SNOSite: exploiting maximal dependence decomposition to identify cysteine S-nitrosylation with substrate site specificity. PLoS One **6**(7), e21849 (2011)
10. Henikoff, S., Henikoff, J.G.: Amino acid substitution matrices from protein blocks. Proc. Natl. Acad. Sci. **89**(22), 10915–10919 (1992). USA
11. Altschul, S.F., et al.: Gapped BLAST and PSI-BLAST: a new generation of protein database search programs. Nucleic Acids Res. **25**(17), 3389–3402 (1997)
12. Pang, C.N., Hayen, A., Wilkins, M.R.: Surface accessibility of protein post-translational modifications. J. Proteome Res. **6**(5), 1833–1845 (2007)
13. Ahmad, S., Gromiha, M.M., Sarai, A.: RVP-net: online prediction of real valued accessible surface area of proteins from single sequences. Bioinformatics **19**(14), 1849–1851 (2003)
14. Ahmad, S., Gromiha, M.M., Sarai, A.: Real value prediction of solvent accessibility from amino acid sequence. Proteins **50**(4), 629–635 (2003)
15. Berman, H.M., et al.: The protein data bank. Nucleic Acids Res. **28**(1), 235–242 (2000)
16. Chang, C.-C., Lin, C.-J.: LIBSVM: a library for support vector machines. ACM Trans. Intell. Syst. Technol. **2**(27), 1–27 (2011)
17. Lu, C.T., et al.: Carboxylator: incorporating solvent-accessible surface area for identifying protein carboxylation sites. J. Comput. Aided Mol. Des. **25**(10), 987–995 (2011)
18. Vacic, V., Iakoucheva, L.M., Radivojac, P.: Two Sample Logo: a graphical representation of the differences between two sets of sequence alignments. Bioinformatics **22**(12), 1536–1537 (2006)
19. Jin, W., Wu, F.: Proteome-wide identification of lysine succinylation in the proteins of tomato (solanum lycopersicum). PLoS One **11**(2), e0147586 (2016)
20. Rardin, M.J., et al.: SIRT5 regulates the mitochondrial lysine succinylome and metabolic networks. Cell Metab. **18**(6), 920–933 (2013)
21. Li, X., et al.: Systematic identification of the lysine succinylation in the protozoan parasite toxoplasma gondii. J. Proteome Res. **13**(12), 6087–6095 (2014)
22. Xie, L., et al.: First succinyl-proteome profiling of extensively drug-resistant mycobacterium tuberculosis revealed involvement of succinylation in cellular physiology. J. Proteome Res. **14**(1), 107–119 (2015)

The Signal Control for the Traffic Network via Image Analysis

Du Dao Huy(✉) and Mui Nguyen Duc

Faculty of Electric Engineering, Thai Nguyen University of Technology, 3/2 street, Tich Luong ward, Thai Nguyen, Vietnam
daohuydu@tnut.edu.vn, nguyenducmui@gmail.com

Abstract. The traffic jam issue is a current important problem which is caused by the rapid increase in population, economic growth and traffic participator. This article represents a signal control method at each intersection to help traffic vehicles to be able to run as fast as possible, and data link among these intersections helps to solve the problem of traffic network in the same time. Thereby, traffic stream management and combination of controlling light signal in traffic network will have a part in reducing traffic jam, saving time and costing for society. A solution for controlling adaptive traffic light signals based on occupied roadway area of vehicles in traffic is presented. The effectiveness of the method is illustrated by an example simulation.

Keywords: Smart traffic · Traffic stream management · Image process · Light signal · Traffic jam

1 Introduction

Currently, most techniques for being used to control traffic lights at each intersection are employing the fixed time. It means that, each traffic stream has fixed time for participators to go or stop at each intersection. The fact that is setting fixed time of each vehicle stream is based on examination about traffic density of each time period or average days. However, traffic streams change irregularly, may increase or decrease unusually, so these are the main limit of traffic light control with fixed preset time. In fact, in populated areas, the dense traffic network combine traffic light control that makes the efficiency of the traffic light system. Currently, a traffic stream is preferential in the traffic network, that is estimating the time of this moving stream to go to the next intersection (based on distance between intersections and average speed allowing) to control the traffic lights of the next intersections [1,4]. However, this control method does not include the sudden increase or decrease of the traffic flow at this stream, so some cases are not suitable. The difficulty when applying smart traffic light control systems must use processors with the high configuration to meet the real-time control. With the system using fuzzy controller [2], they need to have a computer that is difficult to apply and implement in practice. However, today several cities have used the camera system and computer to monitor traffic vehicles. This is the basis for the application of intelligent controllers to control traffic lights [3,5]

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_68

2 The Signal Adaptive Control for Traffic Light Signals at an Intersection

Considering a traffic node in Fig. 1, traffic lights at the opposite branches are controlled equally (the branch 1 is similar to the branch 3, the branch 2 is similar to branch 4), the lights at the adjacent branches have relationships with the others:

$$\begin{cases} t_{R1} = t_{G2} + t_{Y2}; t_{R2} = t_{G1} + t_{Y1} \\ t_{R1} + t_{Y1} + t_{G1} = T; t_{R2} + t_{Y2} + t_{G2} = T \end{cases} \tag{B.1}$$

Where: t_{Ri} is the time of the active red lights, t_{Yj} is the time of the active yellow lights, t_{Gk} is the of the active green lights, T that is the period of controlling traffic. The time for walking at each branch is corresponding with the active red light time at that branch.

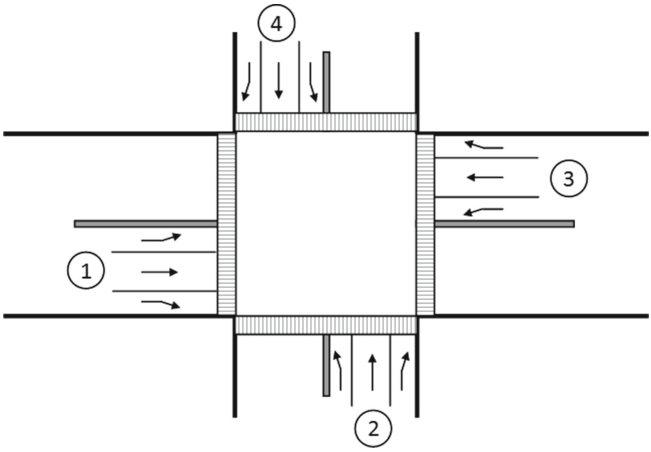


Fig. 1. A traffic light model at a crossroad.

$$\begin{cases} t_{Y1} = t_{Y2} = t_{Y3} = t_{Y4} = const; \\ t_{G1} = t_{G3} = const; t_{R1} = t_{R3} = const; \\ t_{G2} = t_{G4} = const; t_{R2} = t_{R4} = const; \\ t_{R1} + t_{Y1} + t_{G1} = t_{R2} + t_{Y2} + t_{G2} = T = const \end{cases} \tag{B.2}$$

The light control algorithm is simple when there are no vagaries of traffic. This algorithm is not suitable. And in some cases it will cause an obstruction. The idea for this article is to control green lights adapted with the traffic flow at that stream. Red light time of the other stream is corresponding to the change such as (B.1) From that, light controlling period is also changed. Consider the red time that vehicles are waiting for at a branch of a crossroad

shown in Fig. 1, when green light allow vehicles to turn left, go straight, and turn right. Adaptive algorithm will estimate traffic flow at each stream to decide active green light time at that intersection and at the next intersection. In order to do this, the adaptive controller at each intersection will process image at red light time waiting for and estimate green light time (Fig. 2).

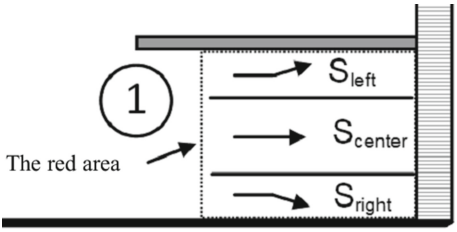


Fig. 2. The area needed to measure at a branch during waiting for red time

Processing image at red light waiting time. Purpose of this step is to calculate the percentage of occupied area of the roadway of vehicles in each stream. Processing image data from camera includes following steps (Fig. 3):
Step1: collecting images at the time when the light is going to switch from red light to green light (yellow light state).



Fig. 3. Image obtained to calculate

Step 2: background subtraction. Markers using to localize: $M_1(x_1, y_1)$, $M_2(x_2, y_2)$, $M_3(x_3, y_3)$, $M_4(x_4, y_4)$ with $y_1 = y_2 = y_{12}$; $y_3 = y_4 = y_{34}$. The calculated area:

$$\sum_{row,col}(x,y) = \begin{cases} y_{12} < y < y_{34} \\ \frac{x_4 - x_1}{y_4 - y_1} < x < \frac{x_3 - x_2}{y_3 - y_2} \end{cases} \tag{B.3}$$

And $\sum_{row,col}(x,y) = 0$ with the others (x,y).

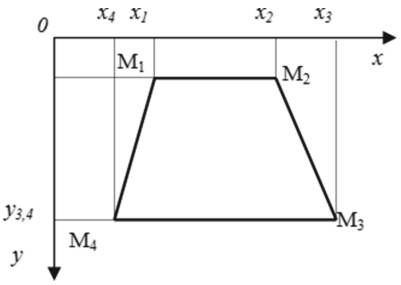


Fig. 4. The localized area

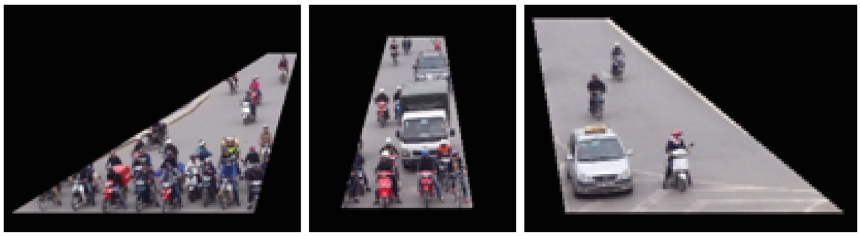


Fig. 5. Images assigned by value o in areas that is not desired

After localizing the area (Fig. 4), we have:

Convert from RGB color space to HSV color space: Values R, G, B divided by 255 values from 0 to 255 convert to range from 0 to 1 (Fig. 5):

$$\begin{aligned} &R' = \tfrac{R}{255}; G' = \tfrac{G}{255}; B' = \tfrac{B}{255}; \\ &C_{\max} = \max(R', G', B'); C_{\min} = \min(R', G', B') \\ &\Delta C = C_{\max} - C_{\min} \end{aligned} \tag{B.4}$$

Hue H:

$$H = \begin{cases} 60^o \times (\frac{G'-B'}{\Delta} \bmod 6) & , C_{\max} = R' \\ 60^o \times (\frac{B'-R'}{\Delta} + 2) & , C_{\max} = G' \\ 60^o \times (\frac{R'-G'}{\Delta} + 4) & , C_{\max} = B' \end{cases} \tag{B.5}$$

Saturation S:

$$S = \begin{cases} 0 & , \quad C_{\max} = 0 \\ \frac{\Delta}{C_{\max}}, & C_{\max} \neq 0 \end{cases} \tag{B.6}$$

Brightness (or Value) V: $V = Cmax$

We have two images having HSV color space, so we compare corresponding points of two images together by using bit wise operation BITXOR. Values of the same pixel of two images will convert to binary value by following rules (Fig. 6):

After performing bit wise operation BITXOR, we convert back to decimal system. The following images in Fig. 7 are results.

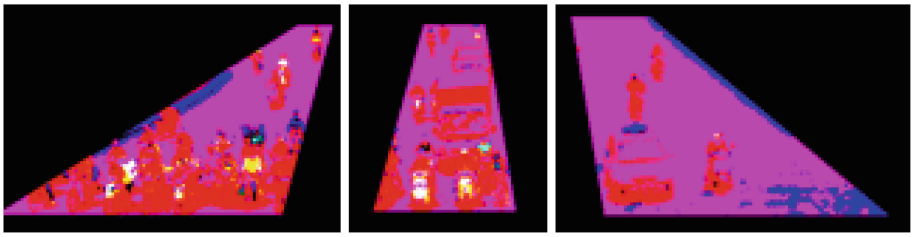


Fig. 6. Image of HSV color space

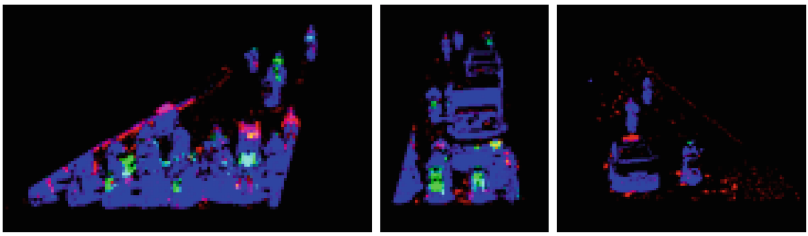


Fig. 7. The images after background subtraction

To calculate occupied area, we need converting to binary color space. HSV color space can not convert to binary color space directly, so we convert from HSV− >RGB− >GRAY− >BINARRY - Converting HSV color space to RGB color space: Assuming that we have HSV color space with $H = [0, 360], S = [0, 1], V = [0, 1]$. Then, we calculate; $H' = \frac{H}{60}$; $X = C(1 - |H' \bmod 2 - 1|)$. And R_1, G_1, B_1 is calculated:

$$(R_1, G_1, B_1) = \begin{cases} (0, 0, 0), & H\, undefined \\ (C, X, O), & 0 \le H' < 1 \\ (X, C, 0), & 1 < H' < 2 \\ (0, C, X), & 2 < H' < 3 \\ (0, X, C), & 3 < H' < 4 \\ (X, 0, C), & 4 < H' < 5 \\ (C, 0, X), & 5 < H' < 6 \end{cases} \tag{B.7}$$

Place $m = V\check{}C$, we have the final result:

$$(R, G, B) = (R_1 + m, G_1 + m, B_1 + m) \tag{B.8}$$

$$\underset{\substack{i:\,1\to row \\ j:\,1\to col}}{Gray[i, j]} = \frac{R_{ij} + G_{ij} + B_{ij}}{3} \tag{B.9}$$

GRAY to BIN:

$$\underset{\substack{i:\,1\to row \\ j:\,1\to col}}{bin[i, j]} = \begin{cases} 1 \text{ if } Gray[i, j] < \text{thresh} \\ 0 \text{ if } Gray[i, j] > \text{thresh} \end{cases} \tag{B.10}$$

Step 3: noise rejection, this is an important step to calculate the area exactly. There are a lot of noise reaction methods such as: Median filter, Bilateral filter, Gaussian filter, lock filter... with this problem. We need removing noise clusters, so we use a Median filter, after we apply Morphological structuring element to perform dilation and erosion of edge of the binary image Step 4: calculating different area. After step 3, we have the binary image with bit 0 (stand for white) and bit 1 (stand for black). To calculate the area, we calculate indirectly through the number of the bit 0s which have in image:

$$the\,bit\,number = \sum_{i=1}^{row}\sum_{j=1}^{col}[bin(i,j)] - \sum_{i=1}^{row}\sum_{j=1}^{col}\left[bin(i,j)|_{bin(i,j)=1}\right] \tag{B.11}$$

When calculating the number of bits, we convert to occupied areas depended on different systems, position of camera, and the distance between the camera and objects.

Estimating green light time.

The occupied areas of traffic vehicles during the time which is waiting for red light at the branch 1 and branch 3 are averaged:

$$\begin{cases} S_{left} = \frac{S_{left_1}+S_{left_3}}{2} \\ S_{center} = \frac{S_{center_1}+S_{center_3}}{2} \\ S_{right} = \frac{S_{right_1}+S_{right_3}}{2} \end{cases} \tag{B.12}$$

The areas in Eq. (B.12) is specific to the need for going at branch 1 and branch 3, hence to estimate time, we need calculating:

$$S = \max\left[S_{left}, S_{center}, S_{right}\right] \tag{B.13}$$

These areas are converted to corresponding time to control green light in the next period:

$$t_{G1}(i+1) = t_{G3}(i+1) = \begin{cases} 5s & S(i) < 5\,\% \\ 10s & 5\,\% \leq S < 25\,\% \\ 15s & 25\,\% \leq S < 55\,\% \\ 20s & 55\,\% \leq S < 75\,\% \\ 25s & 75\,\% \leq S < 100\,\% \end{cases} \tag{B.14}$$

with $t_{G1(i+1)}$ and $t_{G3(i+1)}$ is the time to control green light of branch 1 and branch 3 in the next period. Depending this time, the red lights in the other branch change similarly Eq. (B.15)

$$\begin{cases} t_{Y1} = t_{Y2} = 3s\,(const) \\ t_{R1}(i+1) = t_{G2}(i+1) + t_{Y2}(i+1) = t_{G2}(i+1) + 3 \\ t_{R2}(i+1) = t_{G1}(i+1) + t_{Y1}(i+1) = t_{G1}(i+1) + 3 \end{cases} \tag{B.15}$$

The branch 2 and branch 4 are similar to the branch 1 and branch 3.

3 Adaptive Control for Light Signals in the Traffic Network

Time for controlling green light of branches at node 2 is also separately processed as session 2, however, to be optimal, we control and combine with node 1.

Assuming that traffic nodes are connected together with. The distance between them and average speed allows vehicles to go is given. The branch 1 and the branch 2 of the node 1 are during red light. From the above algorithm, we have determined density of vehicles participating in the branch going straight (S_{center_1}), and time period is to move from the node 1 to the node 2:

$$t_{12} = \frac{l_1 + l_{12}}{v_{TB}} \approx \frac{l_{12}}{v_{TB}} \tag{B.16}$$

with l_1 that is the width of the branch 2 of node 1, l_{12} that is the distance between node 1 and node 2, normally $l_1 << l_{12}$, v_{TB} that is average speed allowed However at node 1, after a period of green light of the branch 1 and branch 2 such Eq. (B.14), traffic vehicles occupy the areas: S_{right_2} and S_{left_4} move to node 2. So period of waiting for the red light at the branch 1 of the node 2 must mention flow of these two branches. The controlling algorithm is shown as (Figs. 8 and 9):

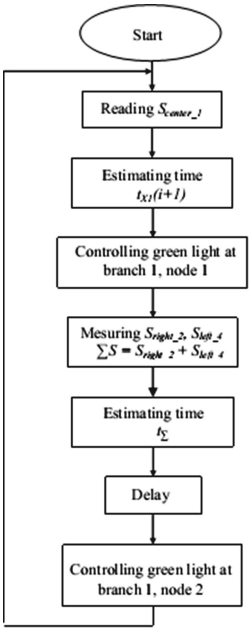


Fig. 8. Controlling flowchart at the node 2

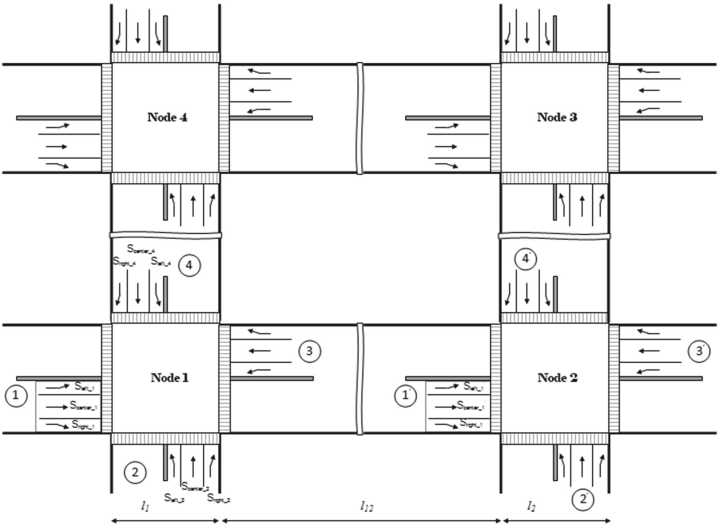
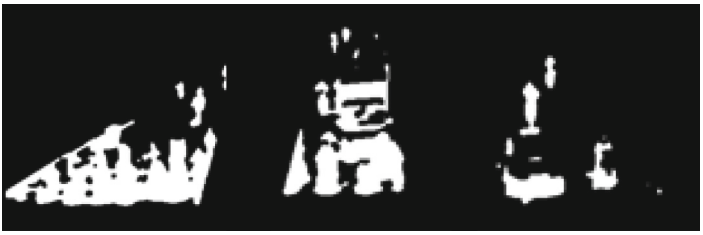


Fig. 9. A traffic network model

Simulation results



(a) pixel number	(b) pixel number	(c) pixel number
1.1126e+05	6.3557e+04	2.3918e+04
S = 80.73 %	S = 43.5 %	S = 17.3 %

With the results from image processing based on Eq. (B.14), We estimate the time for turning the right, going straight, and turning the left are 25 s, 15 s, and 10 s. The time of the next green light is calculated as Eq. (B.13) is 25 s. These results are bases of controlling the traffic light at each node and in the traffic network.

4 Conclusion

The article applies high-tech applications in the signal control system of traffic lights. Using the image processing obtained from the intersection, the occupied areas of roadway in each vehicle lane, since giving the signal control algorithm of traffic lights for lanes based on the occupied areas of roadway of vehicles

are given. The example has demonstrated the effectiveness of the method proposed for traffic network application, since enhancing the efficiency of circulation for vehicles and reducing costs while in traffic.

References

1. Vuong, C.X., Trong, T.V.: The economic influence to select the light cycle at peer-to-peer traffic intersection (2011)
2. Nguyen, C.N.: Research and design the smart traffic light system. Can Tho Univ. J. Sci. (2010)
3. Yousef, K.M., Al-Karaki, J.N., Shatnawi, A.M.: Intelligent traffic light flow control system using wireless sensor networks. J. Inf. Sci. Eng. **26**, 753–768 (2010)
4. Hoai, N.V.: The signal light control follows the green wave. Ho Chi Minh City J. Sci. Appl. (2012)
5. Khattak, M.A.: PLC based intelligent traffic control system. Int. J. Electri. Comput. Sci. IJECS-IJENS (2012)

Toward Cyber-Security Architecture Framework for Developping Countries: An Assessment Model

Nguyen Ai Viet(✉), Le Quang Minh, Doan Huu Hau, Nguyen Ngoc Tuan, Nguyen Nhat Quang, Nguyen Dinh Chinh, Nguyen Van Luc, and Pham Thanh Dat

Information Technology Institute, Vietnam National University, E3 Building, 144 Xuan Thuy, Cau Giay, Hanoi, Vietnam
naviet@vnu.edu.vn

Abstract. This article aims to introduce the cyber security assess model (CSAM), an important component in cyber security architecture framework, especially for the developing country. This architecture framework is built up with the Enterprise Architecture approach and based on the ISO 27001 and ISO 27002. From the holistic perspective based on EGIF developed previously by UNDP group and the main TOGAF features, ITI-GAF is simplified to suit the awareness, capability and improvement readiness of the developing countries. The result of survey and applying in countries as Vietnam, Lao, Myanmar affirms the applicable value of ITI-GAF and the CSAM. The comprehensive, accurate and prompt assessment when applying ITI-CSAM enables the organization to identify the cybersecurity strengths and weaknesses, thereby determine the key parts need invested and its effects to the whole organization's cybersecurity, then build up the action plan for short-term and long-term.

Keywords: Cyber Security Assess Model (CSAM) · ITI-GAF · TOGAF · EGIF

1 Introduction

In recent years, along with the explosive development of Internet infrastructure, smart devices and Internet of Things, information services and social networks, cyber security has become a global real challenge. On one hand, the systems must be flexible and use friendly. On the other hand, it must protect our asset and privacy. In reality, the systems become more and more complex as integrations of many systems deployed by different vendors with different views and interests to cyber security. There must be some architecture to guideline the deployment of information systems while guaranteeing the security. Such an architecture must confront the increasing number of attacks in a variety of forms, tools, environment, at different levels of complexity and severity. It would be a major part of Enterprise Architecture [1,2]. However, in general it is extremely

© Springer International Publishing AG 2017
M. Akagi et al. (eds.), *Advances in Information and Communication Technology*, Advances in Intelligent Systems and Computing 538, DOI 10.1007/978-3-319-49073-1_69

difficult to achieve consensus in Cyber Security. On the other hand, the situation of security is characteristic, as Information System can be designed in a top down approach, while Cyber Security must be designed to adapt to the existing systems. Cyber Security issues are also sensitive to the policy, strategy, top management views and commitments, interpersonal communication. After all, security solutions mainly serve the interests of the organizations, while do not bring new user functionalities, so it is not easy to gain popularity from the beginning. Thus, the popular architecture frameworks like TOGAF, FEA, DODAF,... [3–5] would be too complicated and expensive for Cyber Security. While those tools are superior from the methodological points of view, in practice, it is not easy to implement. Therefore, most architecture frameworks do not cover cyber security issues. To fill this gap, Viet et al. [6] have proposed to apply ITI-GAF [7–12] to construct the Cyber Security Architecture Framework (CSAF) for developing countries. ITI-GAF has an advantage of being simple and easy to adapt to cyber security. In this paper, we will address the assessment model of CSAF. In the implementation process of cyber security projects, the assessment model plays an important role. Firstly, it can be used to enforce the cyber security standards, which are important in the information systems deployed by several different vendors. Secondly, the assessment model can point out the weaknesses in a prioritised order, which help the organizations to prepare an investment and implementation plan to address them. Thirdly, the assessment model can be used to evaluate and monitor the performance of cyber security projects in order to maximize it. In this paper we use the ISO:27001 and ISO:27002 standards to work out the assessment questions. However, this procedure is extendable to adopt other standards as well. We have constructed the assessment schemas with different depths according to various needs of the organizations. Based on these schemas we have designed a web based application to provide assessment services. Although CSAF is constructed for the developing countries, it can be used for more advanced countries as well. The paper is organized as follows: In Sect. 2, an overview of ITI-GAF and the methodology of our work will be presented. In Sect. 3, CSAF will be presented with a strong focus on the assessment model. In Sect. 4, a logical design of a cyber security assessment service based on the CSAF's assessment model will be briefly discussed. Section 5 will discuss the conclusions, learned lessons and future perspectives.

2 Methodology

2.1 Overview of EA and ITI-GAF

EA has been proposed by Zachmann and IBM [1,2] to ensure the interoperability of an information system and to align the business processes, objectives with technology. In 1998, the CIO council and the US' Presidential Budget Bureau have constructed FEA to reduce the failure rate of the US government's IT projects [3]. Soon after that, EA has been built in all advanced countries and became a *de facto* industrial standards, with contributions from more than 350 leading global IT companies and hundreds thousands of projects [4]. ITI-GAF

[6–8] have been developed since 2009 by Nguyen Ai Viet and collaborators at ITI-VNU based on the UNDP's E-GIF [5], TOGAF [4] and other architectures [1–3]. It has been simplified to match the needs and conditions of developing countries. It has been applied successfully in the modeling and design of many important real-life information systems such as E-parliament of Vietnam, 3-level E-office model of Hanoi City and Vietnam's pharmaceutical and cosmetic administration. ITI-GAF is based on an enterprise model consisting of 3 perspectives which are tightly intercorrelated: Operations, Resources and Institutions. The *raison d'être* of an enterprise is its operations, which is supported by suitable resources and institutions. Resources and Institutions must be in accordance with the operations and with each others. For example, if the resources are backward and lag behind the operations, its performance cannot correspond to the competitive business requirements. If the institutions are not updated it might create barriers for the efficient use of resources and effectiveness of the operations.

In our model each perspectives are modeled as consisting of 3 components. The Resources View includes Business Processes, Human Resources and Infrastructures. The Institutions View includes Regulations, Organization and Mechanisms. The Operations View includes External Transactions, Internal Activities and Capability Buildings. The components are also tightly interrelated.



Fig. 1. The ITI enterprise model

With these 3 perspectives, ITI-GAF ensures a fully reflection of all organization's elements and the relationships between them. The interrelations will be use to guarantee the interoperability of an enterprise (Fig. 1). The combination of the three perspectives will bring an overall matrix of 27 intercorrelated and interacting blocks, expressing a holistic view of the organization. The most useful feature of this Enterprise model is that the changes in one block always imply modifications in other blocks accordingly. This feature guarantees the interoperability. For example the infrastructure must satisfy the business needs and should not be over invested to far beyond the skills of the human resources. Organization functionality and responsibility description must enable the currently applied procedures (mechanism) and must be standardized in regulations. The resources and institutions must be developed to support operations efficiently.

All the obstacles and barriers must be removed for the best operational performance.

2.2 Cyber Security Architecture Framework

Today, organizations all over the world are concerned about the assurance of information security. It is worth to note that different strategies must be followed in developed and developing countries. In the former ones, the information systems have achieved a certain level of sophistication and diversity, it is difficult to achieve consensus on a common approach. On the other hand, the awareness of cyber security issues and technical skills in those countries are more advanced. They have passed the first stages to confront with new issues. At the same time, in developing countries [9], the introductions of new technologies and business investments are gradually implemented step by step. The LAN, website and email systems now are the focus. It takes time for them to catch up with their colleagues in developed countries. The same cyber security issues are repeated again and again. Therefore, a cyber security architecture framework should have good benefits. It is also easier to implement CSAF together with the information systems, which are implemented from scratch to avoid the problems in the later stages.

Some organizations, countries applying the Cyber Security framework as NIST [12] to develop Cyber Security. The approach is very expensive, complex and not directly integrated to the enterprise architecture. Therefore, these methods are not suitable for application in developing countries. It is even more complicated if one must build the cyber security measures with the existing EA such as TOGAF, where no CSAF is available at the moment.

As a characteristic aspect of an information system, cyber security is influenced by Operations, Resources and Institutions as well. Since, the regulations will have a stronger influence, the legal framework for cyber security, the habits and the level of people's awareness of cyber security are very different in each country, thus the way these countries face with this issue is very different as well. In that sense, ITI-GAF's generic guidelines turn out to be a very useful and practical tool. In developed countries, basically, infrastructure was invested properly and synchronized; people are accustomed to high-tech services, have sophisticated consciousness of the cyber risks. Therefore the Cyber Security projects can address directly to its objectives. In the developing countries, Cyber Security should be developed based on an architecture framework overarching all aspects of an organization. It must be as simple as possible to implement with an appropriate cost, reduce the learning curves and achieve the consensus easily.

Our recommendation is that if an information system is planned by ITI-GAF together with CSAF, it will maximize its benefits.

3 The Assessment Model

The assessment model based on ITI-GAF should enable organizations to assess the security level of the organizations quickly, accurately and comprehensively.

Through evaluations, each organization will identify the strengths, weaknesses of cyber security in their systems, identify key investment needs and its interactive influence to other parts of the organization, then build up an action plan in the short term and long term to develop the organization and enhance its information security. This is one of the most critical steps for building Cyber Security for organizations. In order to construct the assessment model of CSAF, we use the standards in ISO 27001 [11] and ISO 27002 [12] and classify the measures and requirements according to the ITI-GAF's blocks. ISO 27001 is an international standard for information security management system provides a unified model for establishing, operating, maintaining and improving information security management systems with features such as: risk assessment approach with concentrate on preventative control rather than remedial action, including specifications, application guidelines, requirements, and continuous improvement. ISO 27002 gives guidelines for control practices and implementation of information security for organizations under Sect. 11, 39 control objectives and 133 controls. The projection ISO 27001 and ISO 27002 in the 3*3*3 model provides a comprehensive model which assesses the organization's information security completely, accurately, fast. Depending on the level of detail required, the model can be applied in 3 forms:

- Basic level: applying the basic model with 3 views: Institutions, Resources, and Operations.
- Intermediate level: applying the intermediate model with 9 areas which combine of 3 elements of Institutions (Regulations, Organizations, and Mechanisms) with 3 elements of Resources (Business Processes, Human Resource, and Infrastructure).
- Advance level: applying the advance model with 27 items which combine of 3 elements of Institutions (Regulations, Organizations, and Mechanisms) with 3 elements of Resources (Business Processes, Human Resource, and Infrastructure) and 3 elements of Operations (External transaction, Internal business, and Capability building).

  The assessment criteria are also classified into 4 functions:
- Confidentiality: To prevent the information leaks and unauthorized access to the information and devices.
- Integrity: To ensure that the information are not distorted when being stored or transmitted.
- Availability: To guarantee that the information and devices must be ready to access or use as soon as possible, independent of time and location.
- Non-repudiation: To ensure that the people who access the information or devices cannot deny their actions. The following Fig. 2 show the high level of cooperation between ITI-GAF, ISO 27001, 27003 to build up the questionnaire The result of questions sets gives the basis for a comprehensive review of the organization's cyber security: the strengths, the weaknesses, and correlation between them. Since then the organization can consider critical points need investment and strengthen both in the short term and long term (Table 1).

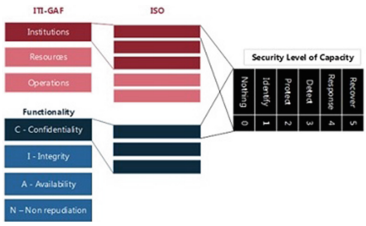


Fig. 2. The questionaire support diagram.

Table 1. Grades of assessment

Grade	Score	Description
None	0	No action implemented
Identify	1	The threats are identified
Protect	2	Measures are implemented to protect the system against the identified threats
Detect	3	Measures are implemented to detect the threats having passed through the protection
Response	4	The system is able to respond efficiently to the detected threats
Recover	5	The system is able to recover from disasters and damages

4 Cyber Security Evaluation Web Service Design

After a period of applying ITI-GAF, the ITI-EA research team has designed an online cyber security evaluation service to help individuals, organizations get more convenient to use the model to assess, and get preliminary understanding on the cyber security. The service is designed as the following figures

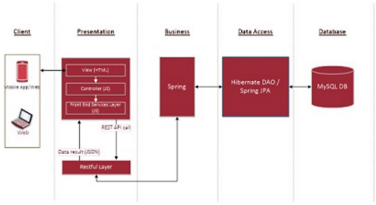


Fig. 3. Logical design for the online CSAF assessment service

Main Functions:

- Give the appropriate question set to the assessment model type (basic, intermediate, or detail) with guideline to answer.
- Give the result of answered question set with some key information
- Ability to reassessment and check improvement progress

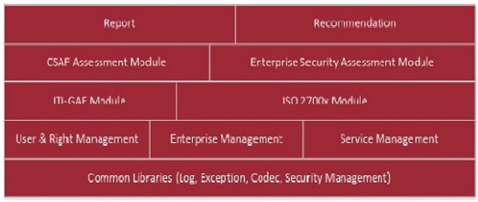


Fig. 4. Business diagram for the online CSAF assessment service

5 Conclusion

CSAF is a guidelines for policy measures, while guaranteeing the operability. It can also maximize the benefits of technology. CSAF based on ITI-GAF is very promising and has been developed to meet those requirements. One hand, it inherits all the good features of the enterprise architecture approach. On the other hand, it has been simplified to match the infrastructure and capacity in the developing countries. The assessment model and the web service designed in this paper can help the organizations, especially in but not limited to developing countries to identify key parts that need improvements. Based on that, it enables those organizations to build up short and long term action plans and to monitor, reassess and adjust the objectives after each development stage. It is the prerequisites for building a whole comprehensive system.

References

1. Zachman, J.A.: A framework for information systems architecture. IBM Syst. J. **26**(3), 276 (1987). IBM Publication G321-5298
2. Zachman, J.A.: Business systems planning, business information control study: a comparison. IBM Syst. J. **21**(3), 31–53 (1982)
3. Chief Information Officer Council: A Practical Guide to Federal Enterprise Architecture, February 2001
4. The Open Group Architectural Framework, TOGAF 9.1 Online Documents. http://pubs.opengroup.org/architecture/togaf9-doc/arch/
5. Viet, N.A. et al.: UNDP's E-GIF (2007)
6. Viet, N.A.: Toward ASEAN-EU Coperation in Cyber Security: An analysis on alignment between EU and ASEAN priorities and objectives - Final Report of CONNECT2SEA project
7. Viet, N.A., Minh, L.Q., Hau, D.H., Lap, N.D., Thuy, D.T.T.: E-organisation assessment based on ITI-GAF. In: Proceeding of FAIR 2014 (2014)
8. Viet, N.A.: 3 Level Model of E-Office for Hanoi City (2012)
9. Nguyen Ai Viet and EA team of ITI: Vietnam E-parliament. Project report (2013)
10. Nguyen Ai Viet and EA team of ITI: Feasibility studies of Vietnam's National Pharmaceutical and Cosmetic Administration's Information System (2012)
11. Hau, D.H.: Research on enterprise architecture building methodologies, propose the methodology for applying in Vietnam. Master Thesis (2014)
12. National Institute of Standards, Technology, Framework for Improving Critical Infrastructure Cybersecurity (2014). http://www.nist.gov/cyberframework/upload/cybersecurity-framework-021214.pdf

Author Index

A
Adreevich, Moldovyan Nikolay, 294
Ahmad, Mohd Ashraf, 238
Akagi, Masato, 6, 490
Akhtar, Pervez, 19
Akif, M. Ahnaf, 197
Alfred, Rayner, 132, 322
Anh, Ngo Hai, 460
Anh, Phan Huy, 510
Anh, Tu Tran, 340
Aziz, Arshad, 19

B
Bich, Diep Nguyen Thi, 572
Bui, Van-Minh, 633
Bui, Xuan, 582

C
Cao, Tuan Dung, 178
Charles, Effiong, 603
Chen, Chan Wai, 347
Chen, Hsi-Min, 367
Chen, Yi-Chung, 367
Chinh, Nguyen Dinh, 652
Chuan, Yoong Yen, 142
Chuyen, Nguyen Anh, 510
Cuong, Duong Chinh, 217
Cuong, Nguyen Duy, 275

D
Dam, Gia Manh, 373
Dao, Thi-Kien, 228
Dao, Toan Thanh, 452
Dao, Van-Lan, 248, 332
Dat, Pham Thanh, 652
Dieu, Nguyen Cong, 153
Dinh, Anh-Tuan, 490
Dinh, Gia Thi, 275
Dinh, Van-Nam, 603
Dinh, Viet Thang, 178
Do, Thi Lien, 94
Do Thanh, Nam, 393
Do Thi, Bac, 393
Dow, Chyi-Ren, 5
Duc, Mui Nguyen, 643
Duong, Hieu N., 84
Duong, Le Khanh, 510

F
Fujita, Hamido, 3

G
Giang, Pham Thanh, 460
Guillet, Fabrice, 302

H
Hadi, Amran Abdul, 238
Hai, Truong Ha, 167
Hao, Nguyen Dang, 275
Haque, Farzana, 470
Hau, Doan Huu, 652
Hien, Nguyen Thi Thu, 312
Hieu, Minh Nguyen, 393
Ho, Bao, 530
Ho, Mau-Viet, 603
Hoa, Nong Thi, 188
Hoang, Anh-Quy, 207, 593
Hoang, Huong-Giang, 421
Hoang, Huy Ngo, 41
Hoang, Van-Phuc, 248, 332
Huong, Nguyen Thi Viet, 113
Husain, Wahidah, 142, 383, 480
Huy, Dao The, 442
Huy, Du Dao, 643
Huynh, Hiep Xuan, 302
Huynh, Hung Huu, 302
Huynh, Nguyen Van, 500

© Springer International Publishing AG 2017

M. Akagi et al. (eds.), *Advances in Information and Communication Technology*,
Advances in Intelligent Systems and Computing, 538,
DOI 10.1007/978-3-319-49073-1

I
Ibrahim, Ag Asri Ag, 322
Ikram, Nassar, 19
Imtiaz, Mehrab, 197

J
Jadin, Mohd Shawal, 238
Jothi, Neesha, 480

K
Kamal, M.D. Rafik, 470
Kazmi, Majida, 19
Khuong, Mai Ngoc, 63

L
Lai, Lai Khac, 255
Le Quang, Minh, 510
Le, Ba Dung, 541
Le, Thi-Lan, 123
Lee, Tzong-Yi, 73
Lei, Tsu-Chiang, 367
Leong, Leow Ching, 132
Linh, Le Hung, 562
Linh, Nguyen Mai, 562
Lu, Dang-Nhac, 357
Luong, Quang Hai, 31
Luong, The Dung, 340

M
Mai, Do Thi, 312
Mai, Luong Chi, 442
Marzuki, Nur Ayuni, 383
Minh, Le Quang, 652
Minh, Nguyen Duy, 312

N
Najib, Muhamad Sharfi, 238
Nakamura, Yasuhiro, 53
Nam, Dao Phuong, 113
Nam, Tran Manh, 500
Nghia, Pham Xuan, 265
Ngo, Thi-Thu-Trang, 357
Ngo, Tung-Son, 603
Nguyen, Duy Phuong, 94
Nguyen, Hai Nam, 53, 294
Nguyen, Hai-Duong, 248
Nguyen, Hai-Minh, 73
Nguyen, Ha-Nam, 357
Nguyen, Hieu Minh, 294
Nguyen, Hoang-Sy, 421, 431
Nguyen, Hong-Tan, 73
Nguyen, Huu-Minh, 612
Nguyen, Ky Minh, 302
Nguyen, Mai, 9
Nguyen, Manh Cuong, 31
Nguyen, Manh Hung, 286, 373
Nguyen, Minh Hieu, 53
Nguyen, Nhat-Tan, 431
Nguyen, Quoc, 530
Nguyen, Tan N., 421
Nguyen, Thang Vu, 452
Nguyen, Thi Hoi, 373
Nguyen, Thi Thu Trang, 178
Nguyen, Thi-Hau, 357
Nguyen, Thu-Trang, 357
Nguyen, Tri Thanh, 104
Nguyen, Trong-The, 228
Nguyen, Tu, 9
Nguyen, Van-Cuong, 603
Nguyen, Van-Danh, 248
Nguyen, Van-Nui, 73, 633
Nguyen, Van-Tinh, 332
Nguyen, Van-Toi, 123
Nguyen, Viet Hai, 541
Nguyen, Xuan Hoai, 104

O
Obit, Joe Henry, 132
On, Chin Kim, 322
Osman, Mohd Azam, 347
Osman, Tousif, 470

P
Pandiyan, Paulraj Murugesa, 322
Pham, Huyen Thanh, 452
Pham, Minh-Trien, 207, 593
Pham, Phuong Thanh, 286
Pham, Tuan, 9
Pham, Van-Bien, 612
Pham-Quoc, Cuong, 530
Phan, Lan Phuong, 302
Phan, Nghia Quoc, 302
Phan, Thanh-Son, 490
Phung, Trung-Nghia, 572
Phuoc, Nguyen Doan, 113
Psyche, Shahreen Shahjahan, 470

Q
Qiu, Sheng-Wei, 367
Quang A, Dang, 403, 413
Quang, Nguyen Nhat, 652
Quang, Vu Vinh, 167
Quy, Ngo Thi Kim, 413
Quynh, Nguyen Thi Thanh, 255

R
Rahman, Rashedur M., 197, 470
Rashid, Nur'Aini Abdul, 480

S
Sainin, Mohd Shamrie, 322
Shadman, Arshiful Islam, 197
Shahiri, Amirah Mohamed, 142, 383
Shin, Kung Ke, 322
Snasel, Vaclav, 84
Son, Le Hoang, 63
Sulaiman, Norizam, 238

T
Talib, Abdullah Zawawi, 347
Tamanna, Fouzia, 470
Tamura, Takumi, 460
Tan, Tran Duc, 31
Tat, Thang Vu, 572
Than, Khoat, 582
Thang, V.V., 622
Thanh, Nguyen Huu, 500
Thi Bui, Anh-Hoa, 431
Thi, Thuong Pham, 104
Thiem, Pham Van, 255
Thinh, Tran Ngoc, 530
Thom, Ho Thi Huong, 228
Towqir, Sheikh Shadab, 197
Tran Pham Thai, Kien, 551
Tran, Cong Manh, 294
Tran, Dinh Hung, 403
Tran, Dinh Que, 286, 373
Tran, Manh Cong, 53
Tran, Phuong T., 421
Tran, Thanh-Hai, 123
Tran, The-Nghiep, 612
Tran, Thi-Xuan, 73
Tran, Truong, 9
Tran, Xuan T., 520
Tri, Nguyen Cao, 84
Trong, Nguyen Thanh, 84
Truong, Quach Xuan, 188
Tu, Trung Nguyen, 41
Tuan, Hung Dao, 294
Tuan, Nguyen Anh, 265
Tuan, Nguyen Ngoc, 652
Tuan, Pham Manh, 393
Tuan, Tran Manh, 63
Tung, Lam Phi, 572

V
Van Chung, Le, 217
Van Hoai, Tran, 84
Van Hoan, Mai, 442
Van Luc, Nguyen, 652
Van Tao, Nguyen, 188
Van Tinh, Nghiem, 153
Van Vinh, Vo, 84
Van, Duc Dang, 41
Van, Thoa Vu, 41
Viet, Nguyen Ai, 652
Voznak, Miroslav, 421, 431
Vu, Dinh Minh, 541
Vu, Duc-Thai, 562
Vu, Tu, 582

W
Wang, Hsin-Ping, 367

Y
Yng, Saw Hui, 480
Yusuf, Nurul Shazrien Amsyha, 238

Z
Zaaba, Zarul Fitri, 347