Michel-Alexandre Cardin
Daniel Krob · Pao Chuen Lui
Yang How Tan · Kristin Wood
Editors

Complex Systems Design & Management Asia

Designing Smart Cities: Proceedings of the First Asia - Pacific Conference on Complex Systems Design & Management, CSD&M Asia 2014

CESAMES
Center of Excellence on Systems Architecture, Management, Economy and Strategy

Springer

Complex Systems Design & Management Asia

Michel-Alexandre Cardin · Daniel Krob
Pao Chuen Lui · Yang How Tan
Kristin Wood
Editors

Complex Systems Design & Management Asia

Designing Smart Cities: Proceedings of the First Asia - Pacific Conference on Complex Systems Design & Management, CSD&M Asia 2014

Springer

Editors
Michel-Alexandre Cardin
Dept. of Industrial & Systems Eng.
National University of Singapore
Singapore
Singapore

Daniel Krob
LIX/DIX
Ecole Polytechnique
Palaiseau Cedex
France

Pao Chuen Lui
National Research Foundation
Singapore
Singapore

Yang How Tan
Defence Science & Technology Agency
Defence Technology Tower A Singapore
Singapore

Kristin Wood
Singapore University of Technology
and Design
Singapore
Singapore

ISBN 978-3-319-12543-5 ISBN 978-3-319-12544-2 (eBook)
DOI 10.1007/978-3-319-12544-2

Library of Congress Control Number: 2014951879

Springer Cham Heidelberg New York Dordrecht London

Printed on acid-free paper

[Springer International Publishing AG Switzerland] is part of Springer Science+Business Media
(www.springer.com)

Foreword

[1] I would like to begin by quoting 2 paragraphs from Aldridge and Augustine "Before exploring the role of systems engineering it is useful to define exactly what is meant by a system. For the purposes herein, a system can be thought of as any collection of two or more entities that interact. Thus, a hydrogen atom fits the definition of a system very neatly — but so, too, does the entire universe. Within those bounds would be the example of a military force — a large collection of interdependent entities which encompasses such diverse functions as training, equipping, maintaining, transporting, communicating, fighting, and much, much more.

To the component designer, an electronic flight control box is a system. But to an airplane designer, an entire aircraft is the system. Yet, to a transportation designer, an airplane is only one element of a much broader system, one which encompasses railroads, ships, automobiles, buses, trucks, bridges and tunnels... not to mention passengers."

The critical components of a complex system are not just the physical components of a system as illustrated by Aldridge and Norman but the invisible interactions between politics, social, economic and environment factors and the many stakeholders of the system.

Since independence in 1965, Singapore has developed many large scale systems The small size of her territory of about 1,400 square kilometers has compelled a tight integration of planning and development of physical systems.

A prime example of a large-scale system is the creation of land by reclamation of swamps in the Western part of the island and in the East from the sea. Lands were needed for industrial development in the Western part of Singapore and for housing development in the Eastern part. The planning of the lands for industrial development was quite different from that for housing. Industrial estates were planned and

---

[1] POINTER - Journals - 2004 - Vol 30 No. 4 - It's a Systems World - After All by E.C. Aldridge and N.R. Augustine. Norman R. Augustine is a graduate of Princeton University and has served as Under Secretary of the U.S. Army, as Chairman and CEO of Lockheed Martin Corporation and as chairman of the National Academy of Engineering. Aldridge and Augustine first worked and wrote together in the 1960s at Douglas Aircraft Company, Santa Monica, California.

developed by JTC – a Government Agency under the Ministry of Trade and Industry. New towns for housing were planned and developed by HDB – a Government Agency under the Ministry of National Development. The complexity of these large-scale projects made large-scale systems engineering a necessity. Our planners and engineers had to learn by doing

The large-scale systems of Singapore can be roughly categorized under three headings, Social Systems, Economic Systems and National Security Systems. The large-scale systems are tightly coupled with changes in one system affecting the others.

Fundamental to Singapore's development is confidence. Our people must be confident that they can create a future for themselves and for future generations. Investors must have confidence that the investments that they will make in Singapore will yield returns without fear of losing their capital. Jobs were created by the economic systems. Social systems provide our people with housing, education and health care. National security is a prerequisite for both economic development and social development. As the nation has very limited resources of people, land and money the allocation of resources is a major function of the Government. Besides balancing current needs, there is also a need to balance short-term with long-term strategic investments

Racial harmony is a pillar for the nation. The second pillar is the tripartite relationship between employers, employees and Government. The integration of the efforts of multiple stakeholders to build a nation is never complete. It is a constant work in progress.

The integration of the Public and Private sectors in the development of large-scale systems is a norm in Singapore. Knowledge and experience developed with each project have helped to build up system engineering practices. These practices reside in each domain of expertise like urban development, public housing, industrial towns, land transportation, marine and offshore industry, aerospace industry, water, energy, environment, parks and defence.

An example of a large-scale system is the water system of Singapore. In the beginning the main effort to deal with water scarcity was in the creation of water catchments and reservoirs and the development of a reliable water supply network. In 2000 with the commercialization of low-pressure membranes, water reclamation of wastewater and desalination of seawater became economically viable. The introduction of NEWater by PUB as a new source of water had helped to create a water industry with exports that benefited our economy.

One of the greatest challenges for large-scale systems engineering is that of integration over time. The "Conception to Retirement" Life Cycle Management System was developed by the engineers in the Ministry of Defence as a solution to the challenge of integration over time. The second challenge is the integration of Operations and Technology. Operations staffs want the new capability for yesterday, using tested and proven technology. Technology staff will want to use technology that will not be obsolete when the project is completed. It is the integration of Operations and Technology in the development of Concept of Operation and Master Plans that will enable projects to achieve the required system operational capability and to sustain it over as long as possible before upgrading or retirement to fulfill the needs of another role.

The longevity of Armoured Fighting Vehicles of the Singapore Armed Forces is an example of how upgrading and changes of role have achieved operational capability needs at the most economic cost.

Systems engineering education began in 1969 with the establishment of the Systems Engineering Group at Singapore University. The first batch of eight master degree students graduated in 1971. The first batch of 39 bachelor degree holders was conferred their degree in 2005. In 2009, Nanyang Technological University (NTU) began offering an MSc degree in Systems Engineering and Project Management. In the same year, a new university - the Singapore University of Technology and Design (SUTD) – with Systems as core was set up in 2009. The first conferment of Bachelor degrees will be in 2015.

The time is right for tighter integration of government agencies, industry and academia to grow the nation.

CSD&M Asia 2014 will provide the opportunity for staff of the three sectors to share and learn from each other. It is equally important that participants will learn about how complex systems are being designed and managed in Europe, North America, Asia and the rest of the world.

It is by the creation of knowledge and sharing of knowledge in complex systems design and management that the goal of achieving quantum leaps in system capability can be achieved.

December 2014

Professor Pao Chuen Lui
Advisor to Singapore Government Agencies
Advisor to President NUS and Advisor to President NTU
Chairman Academic & Research Committee, SUTD

Preface

Introduction

This volume contains the proceedings of the First International Asia-Pacific Conference on "Complex System Design & Management" (CSD&M Asia 2014; see the conference website: `http://www.2014.csdm-asia.net/` for more details).

The CSD&M Asia 2014 conference was jointly organized on December 10-12, 2014 at the National University of Singapore (Singapore) by the three following founding partners:

1. The National University of Singapore (NUS),
2. The Center of Excellence on Systems Architecture, Management, Economy & Strategy (CESAMES),
3. The Ecole Polytechnique – ENSTA ParisTech – Télécom ParisTech – Dassault Aviation – DCNS – DGA – Thales "Engineering of Complex Systems" chair.

The conference benefited of the permanent support of many academic organizations such as Ecole Polytechnique (France), ENSTA ParisTech, National University of Singapore, Singapore University of Technology and Design, and Telecom ParisTech, which were deeply involved in its organization.

Special thanks also goes to Accenture, ADN Singapore, Dassault Aviation, DCNS, Direction Générale de l'Armement (DGA), EDF, IRT SystemX, MEGA International, MIT Press, Sembcorp, Surbana, Thales and The CosmoCompany that were our key industrial & institutional sponsors. The generous support for the Conference Gala Dinner by Sembcorp shall be especially pointed out here.

We are also grateful to ABB, DHI, Gumbooya Pty Ltd., the International Council on Systems Engineering (INCOSE), INCOSE Sector 3 and the INCOSE Singapore Chapter, Land Transport Authority, Ministry of Home Affairs, Veolia Environment and Defence Science & Technology Agency, which strongly supported our communication effort.

All these institutions also helped us a lot through their constant participation to the organizing committee during the one-year preparation of CSD&M Asia 2014.

Many thanks therefore to all of them.

Why a CSD&M Asia Conference?

Mastering complex systems requires an integrated understanding of industrial practices as well as sophisticated theoretical techniques and tools. This explains the creation of an annual *go-between* forum at Asia-Pacific level (which did not exist yet) dedicated both to academic researchers & industrial actors working on complex industrial systems architecture, modeling & engineering. Facilitating their *meeting* was actually for us a *sine qua non* condition in order to nurture and develop in the Asia-Pacific zone the new emerging science of systems.

The purpose of the "Complex Systems Design & Management Asia" (CSD&M Asia) conference is exactly to be such a forum, in order to become, in time, *the* Asia-Pacific academic-industrial conference of reference in the field of complex industrial systems architecture and engineering. This is quite an ambitious objective, that we think possible to achieve, based on the success of the "mother" conference of CSD&M Asia, that is to say the CSD&M conference that organized in France since 2010 with a growing audience (the last 2013 edition grouped almost 300 participants coming from 20 different countries with an almost perfect 50/50 balance between academia and industry).

Our Core Academic – Industrial Dimension

To make the CSD&M Asia conference a convergence point of the academic and industrial communities in complex industrial systems, we based our organization on a principle of *complete parity* between academics and industrialists (see the conference organization sections in the next pages). This principle was first implemented as follows:

- The Program Committee consisted of 50 % academics and 50 % industrialists,
- The Invited Speakers came from numerous professional environments.

The set of activities of the conference followed the same principle. They indeed consist of a mixture of research seminars and experience sharing, academic articles and industrial presentations, software offers presentations, etc. The conference topics cover in the same way the most recent trends in the emerging field of complex systems sciences and practices from an industrial and academic perspective, including the main industrial domains (aeronautic & aerospace, defense & security, electronics & robotics, energy & environment, health & welfare services, media & communications, software & e-services, transportation), scientific and technical topics (systems fundamentals, systems architecture & engineering, systems metrics & quality, systems modeling tools) and system types (transportation systems, embedded systems, software & information systems, systems of systems, artificial ecosystems).

The First Edition of CSD&M Asia in 2014

The CSD&M Asia 2014 edition received 37 submitted papers, out of which the program committee selected 11 regular papers to be published in these proceedings, which corresponds to a 30 % acceptance ratio. Such high selectivity is fundamental for us to guarantee the high quality of the presentations. The program committee also selected 15 papers for a collective presentation during the poster workshop of the conference, that intends to encourage presentation and discussions on other important and emerging issues.

Each submission was assigned to at least two Program Committee members, who carefully reviewed the papers, and in many cases with the help of external referees. These reviews were discussed by the program committee during a meeting held at NUS on 11 July 2014, and via the EasyChair conference management system.

We also chose 13 outstanding invited speakers with various industrial and scientific expertise who gave a series of invited talks covering all the spectrum of the conference, mainly during the two first days of CSD&M Asia 2014. The first and second day of the conference were organized around a common topic – Designing Smart Cities – that gave coherence to all invited talks. The last day was dedicated to a special "thematic session", followed by presentations of all accepted papers as well as a system-focused tutorial in parallel.

Furthermore, we had an Eco-System session in order to provide each participant a good vision on the present status of the systems engineering services and tools offered.

Acknowledgements

We would like finally to thank all members of the Program and Organizing Committees for their time, effort, and contributions to make CSD&M Asia 2014 a top quality conference. A special thank is addressed to the CESAMES (see `http://www.cesames.net/en/`) non-profit organization team and to OPE^3 who helped manage with huge efficiency all the administration, logistics and communication of the CSD&M ASIA 2014 conference.

The organizers of the conference are also greatly grateful to all the following sponsors and partners without whom the CSD&M ASIA 2014 event would simply not exist:

- **Founding Partners**
  - Center of Excellence on Systems Architecture, Management, Economy and Strategy (CESAMES),
  - Ecole Polytechnique - ENSTA ParisTech - Télécom ParisTech - Dassault Aviation – DCNS - DGA - Thales chair "Engineering of Complex Systems",
  - National University of Singapore (NUS).

- **Academic Sponsors**
  - Ecole Polytechnique,
  - ENSTA ParisTech,
  - National University of Singapore (NUS),
  - Singapore University of Technology and Design (SUTD),
  - Telecom ParisTech.

- **Professional Sponsors**
  - Accenture,
  - ADN Singapore,
  - Dassault Aviation,
  - DCNS,
  - Direction Générale de l'Armement (DGA),
  - EDF,
  - IRT System X,
  - Mega International,
  - Sembcorp,
  - Surbana,
  - Thales,
  - The CosmoCompany,
  - The MIT Press.

- **Supporting Partners**
  - ABB,
  - Defence Science & Technology Agency,
  - DHI,
  - Gumbooya Pty Ltd.,
  - International Council on Systems Engineering (INCOSE),
  - INCOSE Sector 3,
  - INCOSE Singapore Chapter,
  - Land Transport Authority,
  - Ministry of Home Affairs,
  - Veolia Environment.

Paris & Singapore, August 20, 2014

Michel-Alexandre Cardin, National University of Singapore, Singapore
Daniel Krob, CESAMES & Ecole Polytechnique, France
Pao Chuen Lui, National Research Foundation, Prime Minister's Office, Singapore
Yang How Tan, Defence Science & Technology Agency, Singapore
Kristin Wood, Singapore University of Technology and Design, Singapore

Conference Organization

Conference Chairs

General Chairs

Daniel Krob, Institute Professor	Ecole Polytechnique, France
Pao Chuen Lui, Advisor	National Research Foundation, Prime Minister’s Office – Singapore

Organizing Committee Chair

Michel-Alexandre Cardin	Assistant Professor, National University of Singapore –Singapore (chair)

Program Committee Chairs

Kristin Wood	Singapore University of Technology and Design – Singapore (academic co-chair)
Yang How Tan	Defence Science & Technology Agency – Singapore (professional co-chair)

Program Committee

The Program Committee consists of 28 members (academic and professional): all are personalities of high international visibility. Their expertise spectrum covers all the conference topics. Its members are in charge of rating the submissions and selecting the best of them for the conference.

Members

Co-Chairs

Kristin Wood	Singapore University of Technology and Design – Singapore (academic co-chair)
Yang How Tan	Defence Science & Technology Agency – Singapore (professional co-chair)

Members

Lynette Cheah	Singapore University of Technoloy and Design – Singapore
Stefano Galelli	Singapore University of Technology and Design – Singapore
Nuno Gil	University of Manchester – United Kingdom
Paulien Herder	TU Delft – Netherlands
Rashmi Jain	Montclair State University – United States
Joseph Kasser	Temasek Defense Systems Institute – Singapore
Nachiko Kohtake	Keio University – Japan
Udo Lindemann	TU Munich – Germany
William Nuttall	Open University – Great Britain
Seiko Shirasaka	Keio University – Japan
Eun Suk Suh	Seoul National University – Korea
Jan Vasbinder	Nanyang Technological University – Singapore
Erik Wilhelm	Singapore University of Technology and Design – Singapore
Akira Yamaguchi	Osaka University – Japan
Maria Yang	MIT – USA
Erik Aslaksen	Gumbooya Pty. – Australia
Abhijit Badwe	ABB – Singapore
Petter Braathen	Memetor – Norway
Hans Eikaas	DHI – Singapore
Kok Seng Foo	Keppel Offshore and Marine Technology Centre – Singapore
Kelvin Ling	Sembcorp – Singapore
Dong Liu	Accenture Technology Labs Beijing – China
Robert May	Veolia Water Asia-Pacific – China
Sin Hin Oh	Land Transport Authority – Singapore
Wilson Oh	Ministry of Homel Affairs – Singapore
Jean-Claude Roussel	AIRBUS Group Innovation – France
Laura Wynter	IBM Research –Singapore

Organizing Committee

The Organizing Committee consists of 18 members (academic and professional) in charge of the logistical organization of the Conference.

Chair

Michel-Alexandre Cardin	Assistant Professor, National University of Singapore - Singapore (chair)

Members

Vladan Babovic	National University of Singapore – Singapore
Saif Benjaafar	Singapore University of Technology and Design – Singapore
Kah Hin Chai	National University of Singapore – Singapore
Guofei Chen	EDF Beijing – China
Aaron Chia	National University of Singapore – Singapore
Mark De Lessio	New York University – United States
Gilles Fleury	Centrale Pekin – China
Stefano Galelli	Singapore University of Technology and Design – Singapore
Liang Kwang Goh	Ministry of Home Affairs – Singapore
Young Jae Jang	Korea Advanced Institute of Science and Technology – Korea
Rashmi Jain	Montclair State University – United States
Edwin Koh	National University of Singapore – Singapore
Harn Wei Kua	National University of Singapore – Singapore
Jianxi Luo	Singapore University of Technology and Design – Singapore
Adam Ng	National University of Singapore – Singapore
Francisco Pereira	Singapore MIT Alliance for Research and Technology – Singapore
Kwong Meng Teo	National University of Singapore – Singapore
Jan Vasbinder	Nanyang Technological University – Singapore

Conference Organization

Invited Speakers

Societal Challenges

- Chong Kheng CHUA, Chief Executive, Land Transport Authority - Singapore
- Daniel HASTINGS, Cecil and Ida Green Education Professor of Engineering Systems and Aeronautics and Astronautics, Massachusetts Institute of Technol-

ogy (MIT) - United States and Chief Executive Officer and Director for SMART, Singapore-MIT Alliance for Research and Technology (SMART) – Singapore
- Pao Chuen LUI, Advisor to Prime Minister's Office, National Research Foundation, Prime Minister's Office – Singapore
- Laura WYNTER, Director, IBM Research -Singapore

Industrial Challenges

- Shiang Long LEE, Executive Director, Institute for Infocomm Research (A*STAR I2R) - Singapore
- David A. LONG, President of the International Council on Systems Engineering (2014 & 2015) INCOSE - United States
- Mun Yuen LEONG, Chief Technology Officer & Senior Director, Infocomm Development Authority of Singapore (IDA) – Singapore
- David TAN, Assistant CEO, Technical & Professional Services Group, Jurong Town Corporation (JTC) – Singapore
- Peng Yam TAN, Chief Executive, Defence Science & Technology Agency (DSTA) – Singapore

Scientific State of the Art

- Saif BENJAAFAR, Head of Pillar and Professor, Engineering Systems and Design, Singapore University of Technology and Design (SUTD) – Singapore
- Richard DE NEUFVILLE, Professor of Engineering Systems and Civil and Environmental Engineering, Massachusetts Institute of Technology (MIT) – United States
- Jeremy WATSON, Professor of Engineering Systems, Vice Dean of Engineering, University College of London (UCL) – United Kingdom

Methodological State of the Art

- Brian COLLINS, Professor of Engineering Policy, University College of London (UCL) – United Kingdom

Tutorial

- Using the INCOSE SE Handbook to Engineer Complex Systems: David Walden, ESEP, Sysnovation, LLC - United States

Contents

Part I: Invited and Regular Papers

Part II: Posters

Part I

Invited and Regular Papers

Model Checking of Energy Consumption Behavior

Shin Nakajima

Abstract. Energy consumption is one of the primary non-functional properties to be addressed early in software system development. Model-based analysis methods are introduced in order to supplement the current practice of runtime profiler techniques. In the present paper, the energy consumption analysis is classified as a duration-bounded cost constraint problem. Specifically, behavioral contracts based on Power Consumption Automata and properties written in terms of weighted linear temporal logic with freeze quantifiers are proposed. In addition, the problem is solved by model-checking of such logic formulas with respect to the automaton.

1 Introduction

Energy consumption is one of the primary non-functional concerns in software-intensive systems, from networked embedded systems to user-centric smartphones, which constitute any modern complex system infrastructure. A software-intensive system, even though functionally correct, may suffer from unexpected energy consumption. Although root causes include design flaws, post-analysis techniques are adapted for checking programs. In Android-based smartphone [1], for example, energy profilers (cf.[20][22]) are used to detect such energy bugs (*e-bugs*) [19]. Since a profiler is a runtime monitor used for checking the energy consumption of running programs, it has the same disadvantages that program testing methods have. Specifically, the profiler can be used only after all of the programs are completed, and the coverage is limited by the supplied test cases.

A model-based energy consumption analysis method is necessary in order to counter the disadvantage of the runtime profiler method. With appropriate design notations, the energy consumption behavior of software-intensive systems can be

---

Shin Nakajima
National Institute of Informatics, Tokyo, Japan
e-mail: `nkjm@nii.ac.jp`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_1

represented and considered early in development. This method is particularly useful for finding design faults caused by, for example, improper use of power management application interfaces of the Android framework.

In the present paper, we propose a model-based formal analysis method for the energy consumption problem, which is formulated as a duration-bounded cost constraint problem. In addition to Power Consumption Automaton (PCA) [13] accounting for the energy consumption behavior, a new variant of linear temporal logic with freeze quantifiers (fWLTL) is introduced so that the duration-bounded cost constraint problem is solved by logic model-checking. For this purpose, we propose a new definition of a PCA in terms of a Timed Automaton with observers [5][6][7]. Previously in [13], the PCA was defined as an n-rate Timed System [4].

The reminder of the present paper is organized as follows. Section 2 describes the concept of the model-based analysis method. Sections 3 and 4 introduce the PCA and fWLTL respectively, and the model-checking method is introduced in Section 5. Related research is discussed in Section 6, and Section 7 concludes the paper.

2 Model-Based Analysis

2.1 *Energy Consumption Behavior*

The energy consumption of a system is attributed to two types of sources. The first is the hardware infrastructure that constitutes the main part of the system including the CPU and memory. These are not controllable by application programs. The second type is peripheral devices that particular application programs use, which include network interfaces, and sensors or actuators. The LCD display, which consumes a lot of energy, also falls into this second category. An application program, running as a backend worker, may not use the display, because it does not provide a GUI for human users.

Although hardware components are direct consumers of energy, application programs that are required in order to make use of these components are responsible for

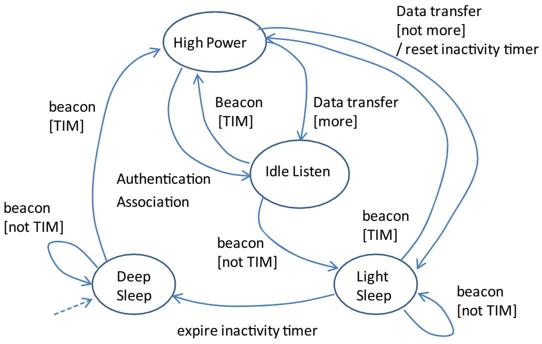


Fig. 1 Example of a PCA Diagram

the consumption. Android-based systems [1], for example, use an aggressive power-saving strategy to make the system sleep when periods of inactivity are detected. Some application programs, however, will make the system awake even when a human user does not touch the screen for some time. The application program calls the power management API of the Android framework such as wake-lock methods, so that the display is kept awake as long as some active locks remain. In the default mode, the wake-lock is maintained by a reference-counting method. An acquired wake-lock must have a matching release call. Otherwise, the wake-lock is kept alive even after the caller program is destroyed. These methods control the energy consumption behavior of the hardware components, and thus the energy consumption is attributed to the caller program. Specifically, improper use of wake-locks result in energy bugs [13][19]. The energy consumption behavior of application programs results from an integrated view of their own functional behavior and the amount of energy consumed by the hardware components.

As an example of energy consumption behavior, Figure 1 shows a state-transition diagram of WiFi client hardware operating in a power-saving mode. The diagram consists of four states, called power states, and several state-transition edges. In Figure 1, `DeepSleep` is the initial state. A transition is enabled to `HighPower` state when a beacon signal, which starts data transfer, is received. The state then alternates between the `IdleListen` and `HighPower` states. In the `IdleListen`, the WiFi client awaits the arrival of a new data frame. When the data frame transfer is over, the `LightSleep` state is entered in order to prepare for a quick restart for data transfer in the near future. An inactivity timer, which causes the transition to the `DeepSleep` state, is also set. The diagram shows that a state-transition sequence consists of many instances of the four power states.

The energy consumption is different for each power state. The `HighPower` state consumes a lot of energy because it needs a large amount of computing resources in order to decode the transferred frames. In the `DeepSleep` state, the electric current is only necessary to activate a small number of hardware circuits and thus the energy consumption is small. If we let $F^j(t)$ be a function of time to represent the rate of consumed energy at a state indexed by j, the total power consumption from time a to b is $P^j = \int_a^b F^j(t)dt$. Since the power states are visited several times as the state transitions progress, the total energy consumption is calculated to be a sum of P^j; $P = \sum_{j=0}^{n} P^j = \sum_{j=0}^{n} \int_{a^j}^{b^j} F^j(t)dt$. We assume here that the state transition sequence is bounded, and that each visited state is indexed by j.

A PCA [13] is a formal model that accounts for the energy consumption behavior. It is informally represented by the diagram in Figure 1. The PCA introduces a linear approximation, so that $dP^j/dt = F^j(t) = M^j$ with a constant M^j for each power state. The constant M^j refers to an average rate of energy consumption at a particular power state j. Then, the total energy consumption P is $\sum_{j=0}^{n} M^j \times (b^j - a^j)$. If we know the constant M^j for each power state, for example, `HighPower` in Figure 1, then the PCA model concisely depicts the energy consumption behavior.

The PCA is regarded as a member of a subclass of Linear Hybrid Automata (LHA) [4] because the dynamics of power consumption variable P^j takes the form of $dP^j/dt = M^j$, and the inactivity timer X is a clock variable of $dX/dt = 1$. However, the PCA is simpler than the LHA in that P^j is an observer and does not have any effect on the state-transition behavior. Contrarily, the clock variable X controls state-transition timings. In Figure 1, for example, the transition from `LightSleep` to `DeepSleep` is enabled by time-out of the inactivity timer. Therefore, a PCA can be regarded as a Timed Automaton (TA) with observers [7], a Weighted Timed Automaton (WTA) [5], or a Priced Timed Automaton (PTA) [6]. These are members of a proper subclass of the LHA and an extension of the TA.

The dynamics in a variant of the PCA are extended to be probabilistic [15]. If we let R be a random variable that is independent and identically distributed (i.i.d.) with respect to probabilistic distribution $f(r)$, the dynamics are written as $dP^j/dt = \alpha(R)\times M^j$ using probabilistic values $\alpha(R)$ with $\alpha\geq 1$. The weight can be used to encode the probabilistic component. Probabilistic or statistical method [21] can be applied to such an extension.

2.2 Energy Consumption Properties

From a naive viewpoint, the properties that must be checked appear simple enough to state that the total amount of consumed energy must be less than a specified maximum value. Since the amount of consumed energy is proportional to the amount of time during which the hardware components are used, the above simple property is eventually violated if the hardware components are used indefinitely. Checking the properties must be limited with respect to time. Thus, the problem involves checking the *duration-bounded cost constraints*, in which the cost refers to the amount of energy consumption. This problem is more general than the duration-bounded reachability of the TA [2]. An expressive and flexible property language is needed to specify precisely and compactly both the durations and constraints.

The example in Figure 1 may be checked to see whether a simple property is satisfied, such that the total amount of energy consumed from `IdleListen` to `LightSleep` is less than a given maximum. A standard linear temporal logic (LTL) formula of the form $\Box(\texttt{IdleListen} \Rightarrow \Diamond\texttt{LightSleep})$ describes the temporal behavior of the desired property. The formula states that *it is always the case that the* `LightSleep` *state is eventually reached if it is in the* `IdleListen` *state*. This LTL formula is extended in order to refer to both the real-time timing and the real-valued energy. To represent the timing, Metric Temporal Logic (MTL) [18] and TPTL [3], for example, have been proposed. However, MTL and TPTL are not applicable to the real-valued energy consumption variables.

We introduce the freeze quantifier that extends the original proposal in TPTL [3] so that the quantifiers can refer to either the energy consumption values or the real-time clocks. If we use freeze quantifiers of the form $\Finv x.\phi^x$, the above-mentioned property to be checked (φ_1) will be written as

$$\begin{aligned}&\Box\Finv^{\tau}x.\ \Finv^{m}u.\ (\texttt{IdleListen}\Rightarrow\\ &\qquad\qquad \Diamond\Finv^{\tau}y.\ \Finv^{m}v.\ (\texttt{LightSleep}\wedge(y\leq x+10)\wedge(v\leq u+50)))\ ,\end{aligned}$$

where x and y refer to the time points, while u and v record the consumed energy at specified states. Here, x and u are *frozen* at the state when the system is in the

`IdleListen` state, and y and v are taken from the `LightSleep` state. If the start and target states of properties to be checked are pre-determined, then the algorithm in [2] or in [5][6] might be used. However, the formula of the type φ_1 may refer to any state proposition to specify the duration. Furthermore, imagine that we have another property φ_2.

$$\Box \text{Ⅎ}^{\tau} x.\, (\texttt{DeepSleep} \Rightarrow\ \Diamond \text{Ⅎ}^{\tau} y.\, (\texttt{HighPower} \wedge (y \leq x + 100)))$$

Although φ_2 is less interesting, we can combine two properties to obtain $\varphi_1 \wedge \varphi_2$ or $\varphi_2 \Rightarrow \varphi_1$ in a flexible manner and describe the expressive property specifications.

If we have such an extension of the LTL with freeze quantifiers, model-based energy consumption analysis is formalized as a model-checking problem. In succeeding sections, we will formally define the PCA and the new extension of the LTL, and then investigate the model checking problem in detail.

3 Behavioral Model

3.1 Power Consumption Automaton

A Power Consumption Automaton (PCA) $\mathscr{A}$ over a set of atomic propositions $Prop$ is defined as a Weighted Timed Automaton (WTA).

$$\langle\, Loc, C, D, \Sigma \cup \{\varepsilon\}, Edg, Flow, Inv, Lab \,\rangle$$

1. Loc is a finite set of locations. Each location corresponds to a power state.
2. C is a finite set of clock variables and D is a set of weight variables. C and D are disjoint ($C \cap D = \emptyset$). For a clock variable x ($\in C$), a constant $n (\in \mathscr{N})$, and an operator $\bowtie \in \{<, \leq, =, \geq, >\}$, constraints of the form $x \bowtie n$ and $x_1 - x_2 \bowtie n$ constitute a set of clock constraints $Z(C)$.
3. Σ is an alphabet that is a finite set of input symbols, and ε is an empty symbol.
4. Edg represents a set of transitions. It is a finite set $Loc \times Z(C) \times \Sigma \times 2^C \times Loc$. The element of Edg, (l_1, g, a, r, l_2), is written as $l_1 \xrightarrow{g,a,r} l_2$, where g is a guard condition in $Z(C)$, a is an input symbol ($\in \Sigma \cup \{\varepsilon\}$), and r refers to a set of clock variables ($\in 2^C$) to reset.

   Furthermore, the PCA is deterministic if the destination location l_2 is determined uniquely at a location l_1 with given a and g.
5. $Flow$ represents the dynamics to account for the change rate of weight variables. For non-negative real $\mathscr{R}_+$ and $\mathscr{R}_+^D$ being $D \rightarrow \mathscr{R}_+$, $Flow : Loc \rightarrow (\mathscr{R}_+^D \rightarrow \mathscr{R}_+^D)$. For a particular weight variable p to keep track of energy consumption and a valuation $w \in \mathscr{R}_+^D$, $Flow(\ell)(w)(p) = dp/dt = M^{\ell}$.
6. Inv is a mapping to clock constraints. $Inv : Loc \rightarrow Z(C)$
7. Lab is a mapping to a set of atomic propositions. $Lab : Loc \rightarrow 2^{Prop}$

The above equation uses valuations for clock variables or weight variables $\nu\in\mathscr{R}_{+}^{C\cup D}$. Reset, delay, and multiplication are defined respectively below.

$$\nu[r](x) \;=\; \begin{cases} 0 & \text{if } x\in r \\ \nu(x) & \textit{otherwise}, \end{cases}$$

$$\begin{aligned} (\nu+d)(x) &= \nu(x)+d, \\ (\nu\times e)(x) &= \nu(x)\times e \end{aligned}$$

For a probabilistic extension of the PCA, $Flow(\ell)(w)(p) = \alpha(R)\times M^{\ell}$, in which α is a probabilistic variable mentioned in Section 2.1.

Finally, two PCAs, where $(C_1\cup D_1)\cap(C_2\cup D_2) = \emptyset$, synchronize on the common set of input symbols $\Sigma_1\cap\Sigma_2$. Whenever PCA_1 makes a discrete transition with a synchronization symbol $a\in(\Sigma_1\cap\Sigma_2)$, PCA_2 also makes a transition.

3.2 Operational Semantics

Semantics of PCA $\mathscr{A}$ is given by a Labeled Transition System (LTS), $\langle\, S,T \,\rangle$. The state space S is a set of tuples consisting of a location l, a clock valuation v, and a weight valuation w. The clock invariants at location l are satisfied.

$$S = \{(l,v,w) \in Loc \times \mathscr{R}_{+}^{C} \times \mathscr{R}_{+}^{D} \;\mid v \models Inv(l)\;\}$$

The transition T consists of regular and stutter transitions; $T = \{\; \xrightarrow{d,e} \;\} \cup \{\; \xrightarrow{d,\varepsilon} \;\}$. The details are explained below.

- Event-trigger discrete transitions $(l_1,v_1,w)\xrightarrow{e}(l_2,v_2,w)$

  $$\exists(l_1 \xrightarrow{g,a,r} l_2) \in Edg \wedge v_1{\models}g \wedge v_2 = v_1[r]$$

- Delayed transitions $(l,v,w_1)\xrightarrow{d}(l,v+d,w_2)$

  $$\begin{aligned} & d \in \mathscr{R}_{+} \wedge w_1 = f(0) \wedge w_2 = f(d) \wedge \\ & \qquad \forall t \in\,]0,d[\; \mid\; v+t \models Inv(l) \wedge df/dt = Flow(l) \end{aligned}$$

  The time advances by an amount of the delay (d) to become $v+d$, and the weight is updated. In the above, $f(t)$ refers to a continuous function differentiable in the open time interval $]0,d[$. For an energy consumption variable p, $Flow(\ell) = dp/dt = M^{\ell}$ at location ℓ, and thus $w_2(p) = M^{\ell}\times d \;+\; w_1(p)$.
- Null transitions $(l_1,v_1,w)\xrightarrow{\varepsilon}(l_2,v_2,w)$

  $$\exists(l_1 \xrightarrow{g,\varepsilon,r} l_2) \in Edg \wedge v_1{\models}g \wedge Flow(l) = \emptyset \wedge v_2 = v_1[r]$$

  The input symbol is empty (ε) and the weights are not changed.
- Regular transitions $(l_1,v_1,w_1)\xrightarrow{d,e}(l_2,v_2,w_2)$

  A delayed transition $(l_1,v_1,w_1)\xrightarrow{d}(l_1,v,w_2)$ followed by an event-trigger discrete transition $(l_1,v,w_2)\xrightarrow{e}(l_2,v_2,w_2)$.

- Stutter transitions $(l, v_1, w_1)\xrightarrow{d,\varepsilon}(l, v_2, w_2)$
  A delayed transition $(l, v_1, w_1)\xrightarrow{d}(l, v, w_2)$ followed by a self-loop null transition at location l, $(l, v, w_2)\xrightarrow{\varepsilon}(l, v_2, w_2)$

A transition sequence is a finite number of $\xrightarrow{d,e}$, or such sequences followed by an infinite number of $\xrightarrow{d,\varepsilon}$. The location l at which the stutter transition is defined is a final state.

Finally, a set of time points $\tau^j \in \mathscr{R}_+$ represents a time progression sequence, where $\tau^0 = 0$ and $\tau^{j+1} = \tau^j + d$ for a delayed transition $\xrightarrow{d}$. For a state $\sigma^j \in S$, a timed point is introduced as $\rho^j = (\sigma^j, \tau^j)$. Then, for a sequence of timed point $\rho = \rho^0\rho^1\cdots$, we have a set of timed sequences $L(\mathscr{A})$ generated by a PCA $\mathscr{A}$, which is written as $\rho \in L(\mathscr{A})$.

4 Property Specification

This section introduces Weighted Linear Temporal Logic with freeze quantifiers (fWLTL) as the property specification language.

4.1 Syntax

The syntax of fWLTL is shown below.

$\pi := c$	Constant $\in \mathscr{N}$ (Natural numbers)
$\mid x + c$	Addition
$\phi := p$	Atomic Proposition $\in Prop$
$\mid \pi_1 \leq \pi_2$	Comparison
$\mid \neg\, \phi$	Logical Negation
$\mid \phi_1 \wedge \phi_2$	Conjunction
$\mid \phi_1 \text{ U } \phi_2$	Until Operator
$\mid \Finv^m x\, .\, \phi^x$	Freeze Quantifier ranging over $C \cup D$
$\mid \Finv^\tau x\, .\, \phi^x$	Freeze Quantifier ranging over time points $\{\tau^j\}$

A variable $x \in Var$, where Var is a countable set of variables, appears free in an fWLTL formula ϕ^x. A closed formula ϕ does not have any free variable. Cost constraints constitute a set $Z(C \cup D)$ because we have $\pi_1 < \pi_2 \equiv \neg(\pi_2 \leq \pi_1)$, and $\pi_1 = \pi_2 \equiv (\pi_1 \leq \pi_2) \wedge (\pi_2 \leq \pi_1)$. Furthermore, the following standard abbreviations are used. $false \equiv p \wedge \neg p$, $true \equiv \neg false$, $\phi_1 \vee \phi_2 \equiv \neg(\neg\phi_1 \wedge \neg\phi_2)$, $\phi_1 \Rightarrow \phi_2 \equiv \neg\phi_1 \vee \phi_2$, $\Diamond\phi \equiv true \text{ U } \phi$ (Eventually Operator), $\Box\phi \equiv \neg(\Diamond\neg\phi)$ (Globally Operator), $\phi_1 \text{ R } \phi_2 \equiv \neg(\neg\phi_1 \text{ U } \neg\phi_2)$ (Release Operator).

4.2 Semantics

We adapt the pointwise semantics for the fWLTL. A number of symbols are introduced.

- A timed state ρ^j=(σ^j, τ^j)=$((l_j, v_j, w_j), \tau^j)$
- A timed state sequence $\rho = \rho^0 \rho^1 \cdots$
- A countable set of variables Var
- An environment $\Gamma \ : \ Var \to \mathscr{R}_+$.
- $\Gamma[x := e]$ assigns x to a value e (e$\in\mathscr{R}_+$) in the environment Γ.

The following satisfiability relations $\models$ defined inductively show the relationship that $\langle\, \rho,\, \Gamma \,\rangle$ satisfies the fWLTL formula ϕ, $\langle\, \rho,\, \Gamma \,\rangle \models \phi$.

$$\begin{array}{ll}
\langle\, \rho^j,\, \Gamma \,\rangle \models\ c & \text{iff}\ \ c \in \mathscr{N} \\
\langle\, \rho^j,\, \Gamma \,\rangle \models\ x + c & \text{iff}\ \ \Gamma(x) + c \in \mathscr{R}_+ \\
\langle\, \rho^j,\, \Gamma \,\rangle \models\ p & \text{iff}\ \ p \in Lab(l_j) \\
\langle\, \rho^j,\, \Gamma \,\rangle \models\ \pi_1 \leq \pi_2 & \text{iff}\ \ \Gamma(\pi_1) \leq \Gamma(\pi_2) \\
\langle\, \rho^j,\, \Gamma \,\rangle \models\ \neg\, \phi & \text{iff}\ \ \langle\, \rho^j,\, \Gamma \,\rangle \not\models \phi \\
\langle\, \rho^j,\, \Gamma \,\rangle \models\ \phi_1 \wedge \phi_2 & \text{iff}\ \ \langle\, \rho^j,\, \Gamma \,\rangle \models \phi_1 \text{ and } \langle\, \rho^j,\, \Gamma \,\rangle \models \phi_2 \\
\langle\, \rho^j,\, \Gamma \,\rangle \models\ \phi_1 \mathrm{\,U\,} \phi_2 & \text{iff}\ \ \langle\, \rho^k,\, \Gamma \,\rangle \models \phi_2 \text{ for some } k \geq j \\
 & \text{and } \langle\, \rho^i,\, \Gamma \,\rangle \models \phi_1 \text{ for all } i\ (j \leq i < k) \\
\langle\, \rho^j,\, \Gamma \,\rangle \models\ \Finv^m x\, .\, \phi^x & \text{iff}\ \ \langle\, \rho^j,\, \Gamma[x := (v^j \cup w^j)(m)] \,\rangle \models\ \phi^x \\
\langle\, \rho^j,\, \Gamma \,\rangle \models\ \Finv^\tau x\, .\, \phi^x & \text{iff}\ \ \langle\, \rho^j,\, \Gamma[x := \tau^j] \,\rangle \models\ \phi^x
\end{array}$$

5 Model Checking

Let $L(\mathscr{A})$ be a set of timed sequences generated by a PCA $\mathscr{A}$. Given a closed fWLTL formula ϕ, the model-checking problem $\mathscr{A}, \Gamma \models \phi$ is defined to ensure $\langle \rho^0, \Gamma_0 \rangle \models \phi$ with an initial empty environment Γ_0 and for all of the timed sequences ρ generated by $\mathscr{A}$ ($\rho \in L(\mathscr{A})$).

We now study the decidability of the above model-checking problem by investigating the existing results on formalisms relating to the PCA and the fWLTL. First, the PCA is a kind of Weighted Timed Automaton (WTA) [5] or Priced Timed Automaton (PTA) [6]. The WTA/PTA are extensions of a Timed Automaton (TA). In the case of the TA, model-checking of the Metric Temporal Logic (MTL) is undecidable, and only model-checking of a fragment of MTL with respect to the TA is decidable [18]. Furthermore, although restricted to considering a subset of the MTL, model-checking of the WTA/PTA is undecidable, but over- and under-approximation techniques are applied to this model-checking [8].

The TPTL [3] is propositional temporal logic with freeze quantifiers, the satisfiability relation of which is defined by timed words generated by the TA, and the freezer quantifier refers to a time point of the binding state (*now*). The TPTL subsumes the MTL [18] as a proper subset, and thus model-checking of the TPTL with respect to the TA is undecidable. However, this model-checkig is decidable when we limit ourselves to the discrete time represented by Natural numbers $\mathscr{N}$.

Based on these existing studies, model-checking of the fWLTL formula with respect to the PCA is undecidable in general. The proposed logic fWLTL is more expressive than the TPTL because the fWLTL can *freeze* weight variables as well as time points. We must introduce some notion of approximations in order to enable analysis. In particular, we will adapt the time-bounded search and maximum time sampling strategy proposed for Real-time Maude [16][17]. Real-time Maude uses *explicit* time model rather than *implicit* or symbolic representation of time. The latter is used in the analysis methods of the TA or the WTA/PTA.

First, the time-bounded search is an under-approximation method in that the search is limited to a finite scope. However, the properties that must be checked are the duration-bounded cost constraints, and thus the time-bounded search seems reasonably effective for finding violations within an appropriate scope. Second, for the case of dense time systems using real-valued time, the state space becomes infinite even if the scope is bounded. Using the maximum time sampling strategy, the sampled points in the timed sequence can be limited to a finite number of points. The strategy is, in a sense, an extension of discrete time sampling where sampling points are selected from $\mathscr{N}$. It chooses sampling points using the maximum possible time advance where sampling points can be $\mathscr{R}_+$. In contrast, the discrete sampling method considers only time points in a constant interval, and thus may miss significant changes between the sampling points. In summary, PCA descriptions are to be translated into Real-time Maude. The translation may follow the method briefly outlined in [14]. This method, however, assumes the PCA to be n-RTS system.

When the PCA has probabilistic weights, model-checking is conducted statistically using the technique commonly known as statistical model-checking (SMC) [21]. Let R be a random variable that is independent and identically distributed (i.i.d.) with respect to a probabilistic distribution $f(r)$. Then, the timed sequences are dependent on the value R_i and thus are represented as $\rho(R_i)$, which is an element of $L(\mathscr{A}(R_i))$. For the fWLTL property ϕ, $B(R_i)$ is defined such that $B(R_i) \equiv (\langle\, \rho(R_i), \Gamma\, \rangle \models \phi)$. Here, $B(R_i)$ follows a Bernoulli distribution $Bin(1,p)$ in which p is the probability that ϕ is satisfied with respect to $\rho(R_i)$. The probability p can be estimated statistically by increasing the number of trials, based on the law of large numbers and the central limit theorem. Therefore, the SMC problem is to generate such a timed sequence $\phi(R_i)$ for R_i and to check $\langle\, \rho(R_i), \Gamma\, \rangle \models \phi$, which can be done using the method of Real-time Maude because the explicit time model explained above is used.

6 Related Work

The present paper discussed a model-based formal analysis method for the energy consumption problem. Specifically, it was formulated as a duration-bounded cost constraint problem. Two formal notations, the Power Consumption Automaton (PCA) and the Linear Temporal Logic with freeze quantifiers (fWLTL) were proposed so that the duration-bounded cost constraint problem was solved by logic model-checking. In the following, we compare the notations of the present study with those of existing research.

The PCA is a kind of Linear Hybrid Automata (LHA) [4]. A PCA can be emulated by an n-rate Timed System (n-RTS) with stopwatch clock variables [13]. Since the energy consumption is observable and does not affect the behavioral specifications, we encode the energy variables as weights in the Weighted Timed Automaton (WTA) [5] or the Priced Timed Automaton (PTA) [6]. The weights in PCA are, however, defined only on states, and not on transition edges. In view of reachability analysis, which is the basis for automatic verification methods, the n-RTA is undecidable while the WTA/PTA, as well as Timed Automaton (TA), are decidable. For the TA, duration-bounded reachability is decidable [2]. Furthermore, optimal or minimum-cost reachability of the WTA/PTA is also shown to be decidable [5][6]. The duration-bounded cost constraints require a property specification language to be flexible enough to express various behavioral aspects. We introduced the fWLTL for expressing properties to check. The problem is solved, in principle, by using the model-checking method. The fWLTL can encode the duration-bounded reachability problem, but does not have any notion of evaluation functions for optimization problems.

Freeze quantifiers were first proposed in the TPTL [3], the semantics of which were defined in terms of timed words generated by the TA. The freeze quantifier of the TPTL refers to a time point of the binding state (*now*). The fWLTL extends this quantifier to the range over weight variables as well as clock variables. The quantified formula $x.\phi^x$ in the TPTL is expressed as $\Finv^{\tau} x.\phi^x$ in the fWLTL. The freeze quantifiers in the Constraint LTL (cLTL) [11] can refer to variables other than clocks, and $\downarrow_{x=m}\phi^x$ in the cLTL is expressed as $\Finv^{m} x.\phi^x$. Finally, Duration Calculus [9] is a formal framework for considering a general class of dense time intervals. Duration Calculus (DC) is very expressive and so the model-checking method is only known for a subclass of DC [12]. We have not yet compared the fWLTL with the DC in detail.

Statistical model-checking of the Linear Priced Timed Automaton (LPTA) is discussed in [10]. The non-deterministic delay in timed transition is replaced by a probabilistic distribution, and the semantics of the LPTA is interpreted to be a stochastic process. Furthermore, the method uses monitor-based model-checking [8], which is based on the explicit time model, rather than an implicit or symbolic representation of time, as in the cases of the TA or the WTA/PTA [4][5][6]. The probabilistic extension of the PCA is weighted probabilistically and thus the role of probability is different from that in the LPTA.

7 Conclusion

In the present paper, we investigated a model-based analysis of the energy consumption problem from the view point of logic model-checking. The problem was formally defined, and it turned out that some approximation was necessary in order to conduct automated analyses. The view presented here is a first step towards the formal model-based analysis of energy consumption behavior. In the future, we intend to investigate in detail a method for using Real-time Maude as the basic verification engine, and to demonstrate the usefulness of the proposed method in detecting energy bugs of Android application programs.

Acknowledgment. The present research was supported in part by JSPS KAKENHI Grant Number 26330095.

References

1. Android, `http://developer.android.com`
2. Alur, R., Courcoubetis, C., Henzinger, T.A.: Computing Accumulated Delays in Real-Time System. In: Courcoubetis, C. (ed.) CAV 1993. LNCS, vol. 697, pp. 181–193. Springer, Heidelberg (1993)
3. Alur, R., Henzinger, T.A.: A Really Temporal Logic. J. Accoc. Comp. Machin. 41(1), 181–204 (1994)
4. Alur, R., Courcoubetis, C., Halbwachs, N., Henzinger, T.A., Ho, P.-H., Nicollin, X., Olivero, A., Sifakis, J., Yovine, S.: The Algorithmic Analysis of Hybrid Systems. Theor. Comp. Sci. 138, 3–24 (1995)
5. Alur, R., La Torre, S., Pappas, G.J.: Optimal Paths in Weighted Timed Automata. In: Di Benedetto, M.D., Sangiovanni-Vincentelli, A. (eds.) HSCC 2001. LNCS, vol. 2034, pp. 49–62. Springer, Heidelberg (2001)
6. Behrmann, G., Fehnker, A., Hune, T., Larsen, K.G., Pettersson, P., Romijn, J.M.T., Vaandrager, F.W.: Minimum-Cost Reachability for Priced Timed Automata. In: Di Benedetto, M.D., Sangiovanni-Vincentelli, A. (eds.) HSCC 2001. LNCS, vol. 2034, pp. 147–161. Springer, Heidelberg (2001)
7. Bouyer, P., Fahrenberg, U., Larsen, K.G., Markey, N.: Timed Automata with Observers under Energy Constraints. In: Proc. HSCC 2010 (2010)
8. Bulychev, P., David, A., Guldstrand Larsen, K., Legay, A., Li, G., Bøgsted Poulsen, D., Stainer, A.: Monitor-Based Statistical Model Checking for Weighted Metric Temporal Logic. In: Bjørner, N., Voronkov, A. (eds.) LPAR-18 2012. LNCS, vol. 7180, pp. 168–182. Springer, Heidelberg (2012)
9. Chaochen, Z., Hoare, C.A.R., Ravn, A.P.: A Calculus of Durations. Information Processing Letters 40, 269–276 (1991)
10. David, A., Larsen, K.G., Legay, A., Mikučionis, M., Poulsen, D.B., van Vliet, J., Wang, Z.: Statistical Model Checking for Networks of Priced Timed Automata. In: Fahrenberg, U., Tripakis, S. (eds.) FORMATS 2011. LNCS, vol. 6919, pp. 80–96. Springer, Heidelberg (2011)
11. Demri, S., Lazić, R., Nowak, D.: On the Freeze Quantifier in Constraint LTL: Decidability and Complexity. Information and Computation 205(1), 2–24 (2007)
12. Meyer, R., Faber, J., Hoenicke, J., Rybalchenko, A.: Model Checking Duration Calculus: a practical approach. Formal Aspects of Computing 20, 481–505 (2008)
13. Nakajima, S.: Model-based Power Consumption Analysis of Smartphone Applications. In: Proc. ACES-MB 2013 (2013)
14. Nakajima, S.: Everlasting Challenges with the OBJ Language Family. In: Iida, S., Meseguer, J., Ogata, K. (eds.) Futatsugi Festschrift. LNCS, vol. 8373, pp. 478–493. Springer, Heidelberg (2014)
15. Nakajima, S.: Platform-Dependence in Model-based Energy Consumption Analysis of Smartphone Applications. NII Technical Report (2014)

16. Ölveczky, P.C., Meseguer, J.: Semantics and Pragmatics of Real-Time Maude. Higher-Order and Symbolic Computation 20(1-2), 161–196 (2007)
17. Ölveczky, P.C., Meseguer, J.: Abstraction and Completeness for Real-Time Maude. ENTCS 176(4), 5–27 (2007)
18. Ouaknine, J., Worrell, J.: Some Recent Results in Metric Temporal Logic. In: Cassez, F., Jard, C. (eds.) FORMATS 2008. LNCS, vol. 5215, pp. 1–13. Springer, Heidelberg (2008)
19. Pathak, A., Hu, Y.C., Zhang, M.: Bootstrapping Energy Debugging on Smartphones: A First Look at Energy Bugs in Mobile Devices. In: Proc. Hotnets 2011 (2011)
20. Pathak, A., Hu, Y.C., Zhang, M.: Where is the energy spent inside my app?: Fine Grained Energy Accoutning on Smartphones with Eprof. In: Proc. EuroSys 2012 (2012)
21. Younes, H.L.S., Kwiatkowska, M., Norman, G., Parker, D.: Numerical vs. Statistical Probablistic Model Checking. J. STTT 8(3), 216–228 (2006)
22. Zhang, L., Gordon, M.S., Dick, R.P., Mao, Z.M., Dinda, P., Yang, L.: ADEL: An Automatic Detector of Engery Leaks for Smartphone Applications. In: Proc. CODES+ISSS 2012 (2012)

UNISENSE: A Unified and Sustainable Sensing and Transport Architecture for Large Scale and Heterogeneous Sensor Networks

Yunye Jin and Hwee Pink Tan

Abstract. In this paper, we propose UNISENSE, a unified and sustainable sensing and transport architecture for large scale and heterogeneous sensor networks. The proposed architecture incorporates seven principal components, namely, application profiling, node architecture, intelligent network design, network management, deep sensing, generalized participatory sensing, and security. We describe the design and implementation for each component. We also present the deployment and performance of the UNISENSE architecture in four practical applications.

1 Introduction

1.1 Background

According to the World Bank, over 52% of the world's population are living in urban areas by the year 2012 [1]. The global trend of increasing urbanization has profound environmental and social impacts. Monitoring environmental status, public infrastructure heath, and commercial service qualify in real time over large scale is critical for urban life quality and sustainable urban development.

1.2 Motivation

Wireless sensor network (WSN) architectures are receiving increasing recognition as a viable approach for urban management. In CommunicAsia 2013, the Infocomm Development Authority of Singapore has revealed the vision of making Singapore 'the first country in the world to be covered by a sensor fabric that connects

Yunye Jin · Hwee Pink Tan
Institute for Infocomm Research, Agency for Science, Technology and Research, Singapore
e-mail: {yjin,hptan}@i2r.a-star.edu.sg

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_2

“everything” and collects data to improve the performance of areas such as healthcare, education, utilities and environmental monitoring’ [2].

Indeed, various WSN deployments have been proposed for practical applications, such as environmental [3] and infrastructure [4] monitoring. The design and implementation of such deployments are highly specific to the particular application of interest, hence not able to incorporate heterogeneous sensing applications. There are also various Internet of Thing (IoT) testbeds that support multiple sensing modalities at various scales, such as MoteLab ([5]), CitySense ([6]), and SmartSantander ([7]). These testbed infrastructures are restricted to specific commercial hardware modules, which limits extension of application range and evolution of technology.

1.3 Contributions

In this paper, we propose UNISENSE, a unified and sustainable sensing and transport architecture for large scale and heterogeneous sensor networks. Our UNISENSE architecture is distinguished from existing WSN architectures and deployments by the following features and advantages.

1. **Heterogeneity:** In UNISENSE, an application profiling (AP) component automatically matches heterogeneous sensing and networking requirements with a combination of suitable hardware parts. The UNISENSE node architecture (UNA) defines a modular approach to interface these hardware parts flexibly. Moreover, we propose a generalized participatory sensing (gPS) component in order to complement the static sensing architecture.
2. **Scalability:** In UNISENSE, we develop a hierarchical network architecture with automatic network discovery mechanisms and multi-hop multi-sink routing protocols. New sensor nodes, network clusters, and third-party legacy systems can be easily incorporated into our network. Our architecture organizes data streams based on sensor applications rather than plain addresses to improve scalability. Back-end servers are designed to dynamically allocate resources on-demand in order to serve thousands of sensor nodes and back-end applications.
3. **Sustainability:** Through the UNISENSE network management (NM) component, system administrators can monitor and visualize the real-time status of the entire network architecture. Moreover, NM provides capability to remotely control and configure network elements, which facilitates sustainable operation. In addition, the security, trust, and privacy (STP) mechanisms of UNISENSE are customized for large scale sensor networks for the authenticity, integrity, and confidentiality of the sensor data.
4. **Deep sensing:** We translate deep learning techniques into the UNISENSE architecture to improve communications, facilitate data recovery, optimize system deployments, and enhance network management. Deep learning uses only simple neuronal operations, which allows sophisticated data processing to be decentralized from the back-end. This feature enables the network to perform more efficiently with its limited resources, function more intelligently in its operations,

and react more quickly to its environment. It also reduces the risk of a single point of failure.

1.4 *Organization of the Paper*

The rest of the paper is organized as follows. Section 2 describes the design of each component of the UNISENSE architecture. Section 3 describes four practical UNISENSE deployments, with some reflections on practical lessons learnt. Section 4 points out the comparative advantages of our UNISENSE architecture compared with other existing WSN architectures. Finally, we conclude the paper in Section 5 and describe the ongoing research and development efforts.

2 The UNISENSE Architecture

2.1 *Overview*

Fig. 1 provides an overview of our UNISENSE architecture. In each application scenario, the AP component characterizes the heterogeneous application requirements and constraints and matches them with an suitable combination of hardware

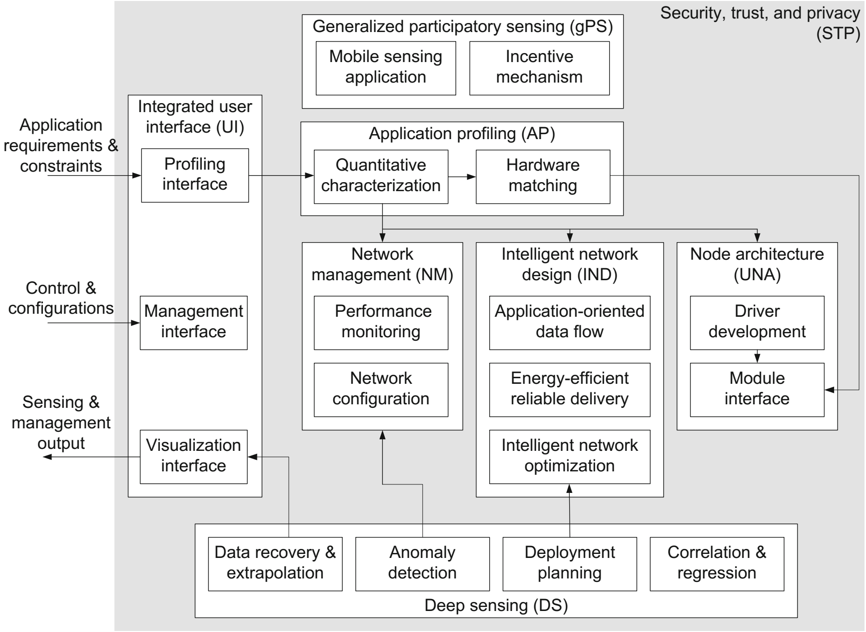


Fig. 1 Overview of the UNISENSE architecture

solutions, which are interfaced by UNA in a modular manner. The IND, NM, and DS components will then optimize the designs and customize the configurations for the specific application requirements and constraints. Clients can get real-time updates of their respective sensing results and control application-specific sensing operations though our UI. The gPS plays an important complementary role to the static WSN infrastructure. The STP component of our architecture ensures each step of the sensing and transportation for each application is secure, trustworthy, and confidential.

2.2 Application Profiling

Application requirements, environmental constraints, and hardware module capabilities are three important aspects concerning the design and deployment of WSNs. In UNISENSE, the AP component quantitatively characterizes these three aspects by specifying the values of an extensive and standard list of parameters in each aspect.

The UNISENSE AP automatically matches a set of specified application requirements and deployment constraints with a suitable set of hardware modules. The matching is performed in an application-oriented manner. Sensor and networking modules are first selected based on sensing and networking requirements. MCU main board is then selected to best interface and power the sensing and networking modules. Lastly, the powering module, including energy harvesters and batteries, are selected based on the overall power consumption of the other selected modules.

2.3 Node Architecture

A UNISENSE sensor node consists of five major components, namely, sensor, actuator, networking, power, and auxiliary. As shown in Fig. 2, the UNISENSE node architecture specifies the standardized interface and schedule mechanism for these components. After AP matches a specific application with the appropriate hardware parts, suitable drivers for each part will be invoked in a modular manner.

The modular nature of our node architecture facilitates collaborative development and technological evolution. Drivers for individual modules can be flexibly

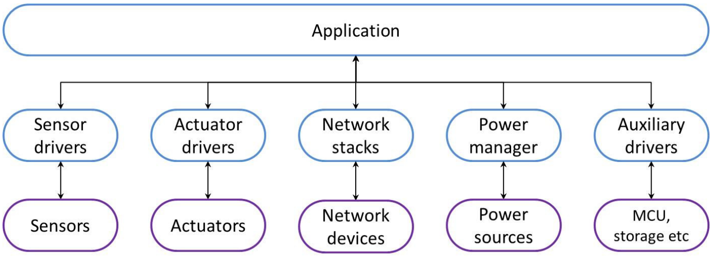


Fig. 2 UNISENSE node architecture model

re-used or modified based on application needs, without affecting other modules' operations. We have also adopted a multi-platform software architecture with the appropriate abstraction models, so that the same software application can be developed for multiple hardware platforms with minimal or no re-development efforts.

We are currently developing an automated device driver invocation system that consistently select software drivers that correspond to the hardware components recommended by AP.

2.4 *Intelligent Network Design*

As shown in Fig. 3, our UNISENSE architecture adopts a hierarchical network organization. Sensor nodes forward their sensor data to the network gateways, which in turn deliver data to the back-end.

Our outdoor sensor network deployments often rely on ambient energy sources, such as solar, which are highly unstable. We have developed an adaptive duty-cycling mechanism that dynamically turns on/off the wireless interface of sensor and gateway nodes in order to achieve energy neutral operation, taking real-time ambient energy availability and networking requirements into consideration.

In order to add sensor nodes to the network architecture in a plug-and-play manner, we have developed an automatic network discovery protocol for newly added sensor nodes to automatically and securely register with nearby gateway nodes and start sensing operations. Moreover, we have also developed a multi-hop multi-sink routing protocol, which enables reliable and fail-safe sensor data delivery.

At the network gateway, we employ a publish-subscribe mechanism, in which gateways "publish" sensor data to the broker (back-end server) and interested users can then "subscribe" to the broker to receive data streams based on topics. This mechanism organizes data based on applications, instead of source-destination address pairs, which facilitates easy integration with third-party legacy systems.

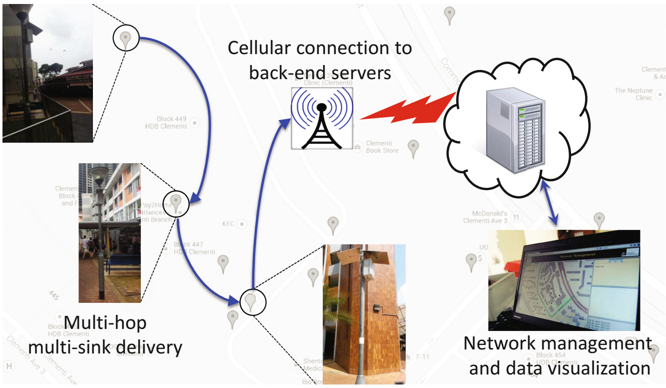


Fig. 3 UNISENSE network architecture

2.5 Network Management and Visualization

In order to ensure sustainable sensing and networking operations of the UNISENSE architecture, the NM component offers a suite of management tools that perform,

1. performance monitoring of sensing, networking, and energy harvesting of the infrastructure at various levels (node, cluster, network) in real time with light weight and low overhead,
2. configuration of the sensing and networking operations and software upgrade of the network infrastructure,
3. mechanisms to track and correlate all changes to the network and its performance for tuning and resource management,
4. alert, diagnosis, and remedy of faults and anomalies,
5. and a unified web-based visualization and control interface for meaningful representation and organization of management information.

2.6 Deep Sensing

The UNISENSE DS component applies state-of-the-art machine learning principles in the design of the WSN architecture by exploiting spatial-temporal correlations in sensor and management data. We employ ideas from a recent advance of neural network technique called *deep learning* [8, 9] to improve the performance and derive actionable insights from the sensor network with limited resources.

By translating deep learning techniques, we have developed the following capabilities for WSN applications.

1. Recovery of missing sensor data due to network loss.
2. Detection of outliers in sensor data as well as anomalies in network operations.
3. Deployment planning in order to fulfill extended coverage requirements with limited hardware resource.
4. Multivariate correlation and regression analysis to assist production and logistic decision makings.

2.7 Generalized Participatory Sensing

The gPS component in UNISENSE exploits sensors on smart mobile devices to complement the static sensing infrastructure.

The major challenges for PS to achieve wide industry adoption are incentive and trustworthiness. In UNISENSE, we have developed an optimized reward scheme that utilizes game theoretics and social networks in order to cultivate a stronger sense of responsibility and accountability among participants and motivate them to contribute higher quality and more trustworthy data.

2.8 Security, Privacy, and Trust

The UNISENSE security solution incorporates a suite of light-weight and energy-efficient cryptographic schemes. It also handles multi-user authentication and intrusion detection. Moreover, it provides a strong security feature into the over-the-air (OTA) programming mechanism for NM purposes.

For both the static and the participatory sensing aspects of the UNISENSE architecture, trust management detects and handles attacks on the trusted network infrastructure. Trust management also assures users of the UNISENSE architecture that their data are handled in an safe and honest way.

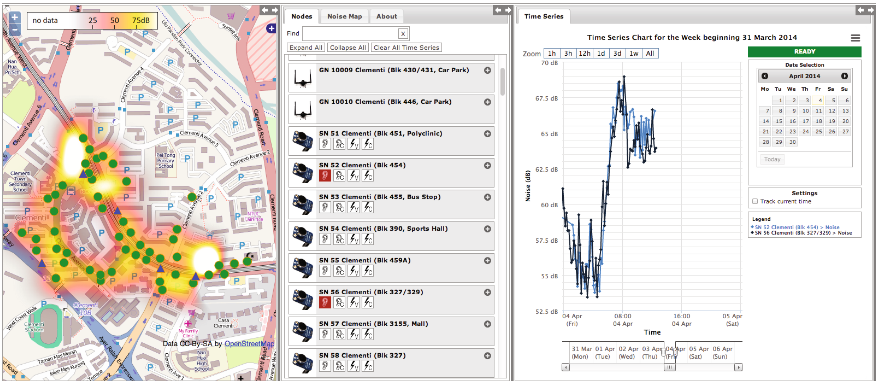
In the UNISENSE architecture, the privacy of the communications is preserved not only at the user end, but also during the transportation in the network by anonymization techniques, improved identity management schemes, and privacy-preserving data analytics and aggregation.

3 Case Studies

3.1 Urban Noise Monitoring

Noise monitoring is an important application in urban management. The conventional practice of ad hoc measurement has poor temporal and spatial coverage. Empowered by the UNISENSE architecture, we have developed and deployed a real-time continuous ambient noise mapping system.

Each noise sensor node consists of low-cost noise sensor, processor, networking, and power hardware, interfaced by UNA. An A-weighting noise level computation algorithm is tailored for the resource-constrained sensor node. The multi-hop



GN 10009 Clementi (Blk 430/431, Car Park)
GN 10010 Clementi (Blk 446, Car Park)
SN 51 Clementi (Blk 451, Polyclinic)
SN 52 Clementi (Blk 454)
SN 53 Clementi (Blk 455, Bus Stop)
SN 54 Clementi (Blk 390, Sports Hall)
SN 55 Clementi (Blk 459A)
SN 56 Clementi (Blk 327/329)
SN 57 Clementi (Blk 3155, Mall)
SN 58 Clementi (Blk 327)

Time Series Chart for the Week beginning 31 March 2014

☐ Track current time

- SN 52 Clementi (Blk 454) > Noise
- SN 56 Clementi (Blk 327/329) > Noise

Visualization for noise map and nodes

Network management and monitor panel

Time series visualization panel for sensor and network performance data

Fig. 4 Integrated back-end interface for urban noise mapping system multi-sink routing protocol delivers data reliably on top of the adaptively duty-cycled sensor and gateway nodes. Battery level and packet delivery ratio are closely monitored through NM toolkits.

The DS component performs missing data recovery and continuous noise map construction at the back-end in real time. As shown in Fig. 4, we have created an integrated interface that visualizes the continuous noise map and time-series. Node operation status can also be conveniently monitored for network management purposes through the same interface.

By Mar. 2014, we have successfully deployed two noise mapping testbeds in two typical urban sites in Singapore, namely, Clementi Town Central and Jurong Lake District. Each testbed consists of 50 solar-powered noise sensor nodes and 4 solar-powered gateway nodes installed on the lamp posts. Every five minutes, each sensor node measures and updates noise level, in dBA, to the back-end server, which in turn updates the time-series and spatial heat map visualizations.

3.2 Public Infrastructure Monitoring (Smart Bins)

Currently, most rubbish bin cleaning service are conducted in a routine manner. Resources are often wasted on clearing some half-filled rubbish bins, while having other overflowing. The Smart Bins deployment aims to solve this problem by providing a real-time update of the bin fullness level, allowing operators to deploy resources in an adaptive manner rather than a routine manner.

Leveraging on the UNISENSE architecture, Smartbins utilizes ultrasonic sensors to sense the fullness level of the bin, and update the back-end through the hierarchical network infrastructure. We also use GPS and accelerometer modules to monitor the location of the bins as they are susceptible to movements by the public, as well as provide accurate timestamps for data. Fig. 5 shows a sensor node for smart bins.

By Apr. 2014, we have deployed 11 sensors, 4 multi-hop and 7 single-hop, over Geylang Road and Guillemard Road in Singapore. With an update interval of five

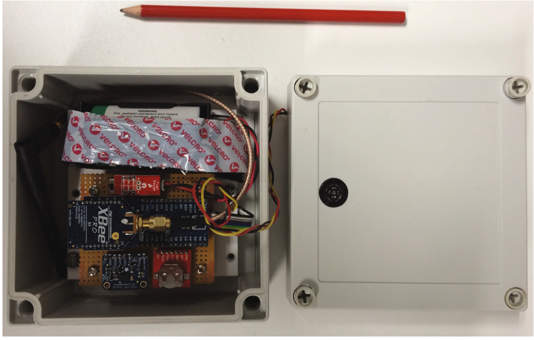


Fig. 5 A Smart Bins sensor node

minutes and a 13 Ah non-rechargeable battery, none of the sensors have required battery replacements since our deployment commenced in Dec. 2013. We are able to achieve more than 90% packet delivery ratio for all the 4 bins with multi-hop routing scheme.

3.3 Event Management (CuteBit)

3.3.1 Application Scenario

CuteBit is a real-time event participation system empowered by the UNISENSE architecture. It provides an intuitive button-less UI to support voting, ranking, and participatory games by simply flipping the CuteBit to different faces.

Each CuteBit is a battery-powered MCU main board equipped with an accelerometer and a 802.15.4 wireless module. The user inputs are captured by sensing the orientation of the CuteBit. A multi-sink routing protocol delivers user input reliably to the gateway nodes in real time.

On 8 May 2013, 42 CuteBits were deployed for a dinner event. One CuteBit sensor node was placed on each dinner table for participation input from this table. The CuteBit system performed three tasks during the event successfully, namely, collaborative event launch, voting, and crowd activity level monitoring. Fig. 6 is a screenshot of the voting visualization.

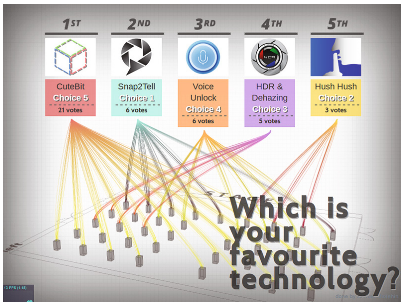


Fig. 6 Voting visualization interface for CuteBit

3.4 Vertical Farming Monitoring

Monitoring vertical farming is a new application domain for WSNs. By deploying WSN on vertical farm towers, we aim to, (i) increase the crop yield through in-depth sense-making of the sensor data and crop growth bio-data, and (ii) improve productivity through reduced reliance on manual monitoring.

We deployed a total of 115 solar powered sensor nodes on a vertical farm to measure multiple environmental parameters including ambient light, temperature, and humidity. The sensors were deployed on 115 rotating towers inside greenhouses for a period of 30 days. All sensor nodes are adaptively duty-cycled to achieve energy neutral operation using solar power. As a result, all sensor nodes have achieved 100% uptime over the 30 days deployment. The DS component performed correlation and regression analytics on the multi-modal sensor data with crop growth bio-data. Fig. 7 show the multi-modal sensing node.

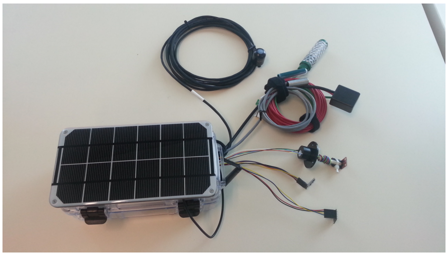


Fig. 7 Multi-modal sensor node for farm use

3.5 Practical Issues and Lessons Learnt

The design, implementation, and deployment of the four testbeds have provided us valuable experience for practical WSN realizations. We summarize them as follows.

1. In practical environments, many dynamic factors, such as weather conditions and traffic flows, significantly affect the operating conditions of the WSN deployment. Therefore, it is advisable to always over-provision for energy harvesting capabilities and wireless antenna gains in order to ensure network uptime and connectivity, respectively.
2. Operating condition in many deployment environment can be hostile to the node hardware. Weatherproof and insulation efforts are the keys to protect sensor node hardware from corrosion or short circuit. Programs and scripts running in sensor and gateway nodes should always be designed with fail-safe and auto-recovery features.
3. At both the node and the network level, our architecture design adopts a modular approach. This not only enables future re-use of both hardware and software modules, but also makes development collaboration, debugging, and upgrading much easier. In the setup and maintenance of our deployments, we have taken full advantage of the modular design of the nodes, replacing and upgrading the parts in a plug-and-play manner.

4 Related Work and Competitive Advantages

Various WSN deployments have been utilized for practical applications. In [3], a WSN was deployed to monitor volcano activities. In [4], a WSN was deployed to manage a combined sewage overflow control system. In these cases, the network planning and the hardware choice are highly specific and fixed to the particular application requirements. Such an application-specific approach cannot support the coexistence of heterogeneous sensing applications in a cost-effective manner. In contrast, our UNISENSE architecture utilizes the AP component to quantitatively characterize the application requirements and constraints, while UNA allows the most suitable hardware modules for each application to be flexibly interfaced together.

Various WSN and Internet of Thing (IoT) testbeds can support multiple sensing modalities and applications. For example, MoteLab [5] is an indoor WSN testbed, with dedicated energy measurement, reprogramming, and management interfaces. CitySense [6] is an outdoor urban WSN testbed consisting of embedded Linux PCs with weather and air quality sensors. The SmartSantander [7] testbed consists of $20,000$ heterogeneous IoT sensors and devices. Although featured with multiple sensing modalities and reprogrammability, these testbeds are restricted by specific commercial hardware modules, which limits the long term technical evolution. Moreover, such testbeds usually incorporate a dedicated management and evaluation sub-infrastructure, which incurs high hardware cost. In contrast, the UNISENSE NM component manages the network infrastructure in an in-network manner with minimum extra hardware and overhead.

To the best of our knowledge, the UNISENSE architecture is the first to have a dedicated machine learning component (DS). The learning capability enables outlier and anomaly detection, missing data recovery, and mining for actionable insights, which brings our infrastructure closer to real world applications.

5 Conclusion and Ongoing Work

In this paper, we have introduced the UNISENSE architecture for large scale heterogeneous WSNs, as a technology enabler for smart cities realization and management. We described the design approach and implementation progress for each principal component of the architecture. We have also presented the deployment and performance of four practical WSNs supported by our architecture.

We point out future work in the following aspects. The UNA will be expanded to support more hardware platforms and interface types in our architecture. Besides the existing automatic hardware matching capability, we aim to automate the choice of driver software in the AP component. For network planning, an IND toolkit will be designed to intelligently determine network topologies, select suitable protocols, and configure the optimal operating parameters.

A suite of scalable NM methods for configuration changes logging, performance tracking, and higher-order fault detection will be developed. An integrated security, trust, and privacy solution suite will be developed for resource-constrained WSN nodes. Deep learning techniques will be translated to facilitate distributed processing, simplified network codec, response time improvement, and data compression. For gPS, we will focus on the design and enhancement of incentive schemes as well as user data quality evaluation mechanism.

References

1. The World Bank, Data on Urban Population (July 2013), `http://data.worldbank.org/indicator/SP.URB.TOTL.IN.ZS/countries/1W?display=graph`
2. CommunicAsia 2013 The Official Daily News, IDA Envisions Sensor Fabric for Singapore (June 2013), `http://www.communicasia.com/post-show/downloads/CMMA2013_showdaily03.pdf`
3. Werner-Allen, G., Lorincz, K., Ruiz, M., Marcillo, O., Johnson, J., Lees, J., Welsh, M.: Deploying A Wireless Sensor Network on An Active Volcano. IEEE Internet Computing 10(2), 18–25 (2006)
4. Montestruque, L., Lemmon, M.: CSOnet: A Metropolitan Scale Wireless Sensor-Actuator Network. In: Proc. International Workshop Mobile Device and Urban Sensing (2008)
5. Werner-Allen, G., Swieskowski, P., Welsh, M.: MoteLab: A Wireless Sensor Network Testbed. In: Proc. IEEE IPSN, pp. 483–488 (2005)
6. Murty, R.N., Mainland, G., Rose, I., Chowdhury, A.R., Gosain, A., Bers, J., Welsh, M.: CitySense: An Urban-Scale Wireless Sensor Network and Testbed. In: Proc. IEEE Conference on Technologies for Homeland Security, pp. 583–588 (May 2008)
7. Gluhak, A., Krco, S., Nati, M., Pfisterer, D., Mitton, N., Razafindralambo, T.: A Survey on Facilities for Experimental Internet of Things Research. IEEE Communications Magazine 49(11), 58–67 (2011)
8. Hinton, G.E., Salakhutdinov, R.R.: Reducing the Dimensionality of Data with Neural Networks. Science 313(5786), 504–507 (2006)
9. Lee, H., Grosse, R., Ranganath, R., Ng, A.Y.: Convolutional Deep Belief Networks for Scalable Unsupervised Learning of Hierarchical Representations. In: Proc. Annual International Conference on Machine Learning, pp. 609–616. ACM (2009)

Black-Box Optimization of Lighting Simulation in Architectural Design*

Alberto Costa**, Giacomo Nannicini, Thomas Schroepfer, and Thomas Wortmann

Abstract. This paper deals with an application of optimization in architectural design. Formally, we consider the problem of optimizing a function that can only be evaluated through an expensive oracle. We assume that the analytical expression of the function is unknown and first-order information is not available. This situation frequently occurs when each function evaluation relies on the output of a complex and time-consuming simulation. In the literature, this is called a *black-box* optimization problem with costly evaluation. This paper presents a black-box problem from architectural design: we aim to find the values of the design variables that yield optimal lighting conditions inside a building. The building façade is described as a parametric model whose parameters are the design variables. We tackle this problem by adapting the Radial Basis Function (RBF) method originally proposed by Gutmann (2001). Experiments indicate that our open-source implementation is competitive with commercial software for black-box optimization, and that it can be a valuable decision-support tool for complex problems requiring time-consuming simulations. The usefulness of this approach goes beyond the specific application in architectural design.

1 Introduction

We consider the problem of optimizing an unknown function given as an oracle. We assume that the oracle is expensive to evaluate, so that estimating partial derivatives

---

Alberto Costa, Giacomo Nannicini, Thomas Schroepfer, Thomas Wortmann
Singapore University of Technology and Design, 138682 Singapore
e-mail: {`costa,nannicini,thomas.schroepfer`}`@sutd.edu.sg`,
`thomas_wortmann@mymail.sutd.edu.sg`

* Supported by IDC grant IDG21300102.

** Corresponding author.

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_3

by finite differences is impractical. In the literature, such a problem is known as a *black-box* optimization problem with costly evaluation.

Black-box optimization finds many applications. Our main motivation stems from architectural design. In recent years, simulation has been increasingly employed as a decision-support tool in this field. As a case study, we optimize a building façade in terms of daylighting and glare to achieve a more sustainable design. The façade is described as a parametric model whose parameters are the decision variables. Daylight simulation that is based on site-specific climatic data can, over a prescribed period, predict brightness inside the building as well as glare. This quantitative information serves as a performance measure, i.e., an objective function. We assume that the desired characteristics can be combined into a single objective function that may include multiple terms, for instance penalties for violating constraints based on the simulated visual comfort of building's occupants. Our goal is to optimize this objective function, and thus to find optimal values for the parametric model of the design. However, each run of the simulation takes considerable time, depending on the desired accuracy. Thus, we want to find values of the decision variables (design parameters) that achieve a close to optimal value for the objective function, while performing only a small number of function evaluations to limit computing times.

Current architectural design practice often relies on experience and best practices rather than on optimization techniques. Better-informed decisions in the design stage of a building can lead to e.g. better environmental performance. Reliable solutions for these types of design problems are crucial for the development of sustainable architecture and highly liveable future cities. Even if optimum efficiency may not always be attainable in practice, optimization techniques provide invaluable information to the architect and thus create new design possibilities. One reason for today's absence of optimization techniques in the design process is the time-intensity of simulations. This constraint is exacerbated by the type of optimization algorithms that are usually proposed for architectural design, that is, genetic and swarm-based algorithms, e.g., [3, 17, 18]. Such algorithms converge on a solution only after performing many simulation runs. The RBF method described in this paper aims to provide much faster convergence, while offering flexibility for designers in defining the performance measure, i.e., the objective function.

A common approach to black-box problems is to use several evaluations of the oracle to build a model of the objective function. This model is also called *response surface* or *surrogate* model. Examples of this approach are the Radial Basis Function (RBF) method of [6] (see also [19]) and the kriging-based Efficient Global Optimization method (EGO) of [11]. Despite the appealing theoretical properties of EGO, empirical evidence suggests that the RBF method is more effective on engineering problems [9]. The RBF method has many potential applications in engineering design beyond the case study presented here, in particular for optimizing the performance of complex physical devices such as engines, see e.g., [1, 7].

The rest of this paper is organized as follows. Section 2 presents a case study of optimal decision-making for complex problems in architectural design. In particular, we introduce a method to provide mathematical models to assess daylighting quality based on the designers preferences. Section 3 provides an overview of the RBF method for black-box optimization, discusses possible extensions, and describes our open-source implementation. A brief computational evaluation shows that our implementation is competitive with commercial applications from engineering design, and orders of magnitude more efficient than genetic algorithms, which are commonly used in architectural design.

2 Application to Architectural Design

The conflicting objectives in architectural design and complexity of the simulations mandate a black-box approach to architectural optimization. Accordingly, the literature on optimization in the architectural design is dominated by discussions of genetic algorithms, e.g., [13, 15]. Other approaches to optimization in the architectural field, such as particle swarm and ant colony optimization, usually are swarm-based, e.g., [12]. Genetic algorithms are also the most likely technique to be employed by architecture practitioners and students. Grasshopper, a free plug-in for the popular 3D-modelling application Rhinoceros™, includes an easy-to-use evolutionary solver named Galapagos. However, in practice, evolutionary and swarm-based approaches require the evaluation of many design candidates via simulation, which likely is one of the reasons why their application in design practice has remained limited. The RBF method presented in this paper conserves the advantages of a black-box approach while greatly reducing the required number of design candidate evaluations, see Section 3.4.

Several aspects make the RBF method especially promising for applications in architectural design. First, being a black-box method, it requires no knowledge about the internal structure of the optimization problem. Although typically in the literature only specific aspects of a building design are optimized, even these smaller problems are characterised by multiple, competing objectives. For example, in optimizing a load-bearing structure, the designer often aims at minimizing both structural displacement and weight, see [12, 13]. Similarly, in façade design, opening sizes have to balance daylight and shading requirements to avoid heat gains and glare. The case study presented in this paper is located in a tropical climate, and aims at maximizing daylight while minimizing glare. Combining these objectives with the complex physics being simulated, we obtain an objective function such that the relationship between inputs and outputs is not analytically available.

The RBF method has two other important benefits over evolutionary or swarm algorithms that are especially important for architectural design. First, the surrogate model created by the RBF method approximates the design space implied by the parametric model and the performance criteria. The surrogate model thus allows the approximate evaluation of different design scenarios around the optimum without additional simulations. Design theorist Roy Woodbury conceives of architectural design as design space exploration and highlights the need for computational tools that support this exploration [23]. From the perspective of design as exploration, the RBF method is a tool to approximately but quickly evaluate the sensitivity of the performance measure with respect to the design parameters. A discussion on how to compute a confidence level of this sensitivity analysis is given in Section 3.3.1. Second, one can use the RBF method with different objective functions without repeating time-intensive simulations. This is because in many situations, such as daylight simulation, one can save the (complex) output data of a simulation run to allow the evaluation of a different objective function relying on the already available output data. In this way, we quickly obtain a model of the modified performance criterion, and can hot-start the optimization method to converge much faster than restarting from scratch. These two aspects at least partially address an important criticism of optimization methods in the architectural design field: optimization has to rely on pre-defined performance criteria and is thus difficult to integrate into complex design processes with shifting performance criteria. This difficulty was succinctly formulated by design theorist Horst Rittel in his famous chapter “Planning Problems are wicked Problems” [20]:

> The methods of Operations Research [. . .] become operational, however, only after the most important decisions have already been made, i.e., after the problem has already been tamed. Take an optimization model. Here the inputs needed include the definition of the solution space, the system of constraints, and the performance measure as a function of the planning and contextual variables. But setting up and constraining the solution space and constructing the measure of performance is the wicked part of the problem. Very likely it is more essential than the remaining steps of searching for a solution which is optimal relative to the measure of performance and the constraint system.

The RBF method attempts to alleviate the “wickedness” of architectural design problems by allowing the approximate exploration of the design space, and, perhaps more importantly, the changing of performance criteria with significantly reduced effort.

As a case study for the integration of the RBF method into architectural design processes, the method is employed in the design of a mixed-use high-density church building in Singapore. The church is designed with a performative façade that modulates daylighting conditions on the interior due to the differentiated inclinations of its disk-shaped louvers, see Figure 1. The value of the angles of the small, disk-shaped louvers are the design variables. We associate one design variable with the opening angle of a group of louvers, instead of individual louvers. This simplification serves three purposes: it reduces the dimension of the optimization problem, guarantees a more unified visual effect of the façade, and standardizes the façade elements for easier construction. We further reduce the optimization problem complexity with two techniques from systems architecture, namely decomposition and integration [14]. For an individual room inside the building, lighting conditions depend only on the part of the façade adjacent to that room. Hence, we can decompose the problem by considering each room separately, and find the optimal configuration of the corresponding disk-shaped louvers using the RBF method. Finally, we integrate these partial solutions to obtain a global optimal configuration for the façade as a whole. Accordingly, in the following we consider only one individual room.

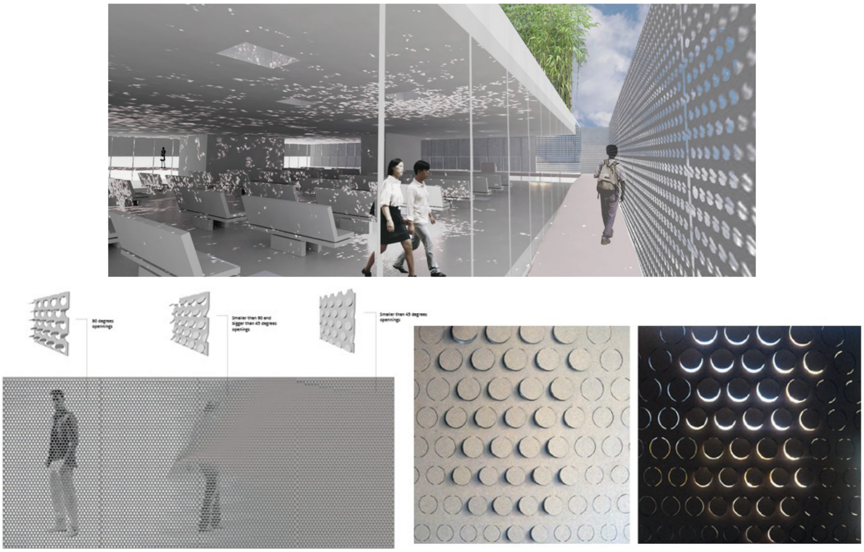


Fig. 1 Façade of the church building with disk-shaped louvers

Each of our design variables represents the opening angle of a group of louvers, and takes values between 0° (louvers closed) and 180° (louvers fully open). Daylighting is simulated with DIVA, a plugin for the 3D-modelling software Rhinoceros™ 3D [10]. DIVA yields illuminance values inside the building. Typically, running such a daylighting simulation is very time consuming and may take up to several hours. To define an optimization problem we need an objective function. In other words, given an output of the lighting simulator, how can we assess its quality? The quality of given lighting conditions is assessed in terms of two metrics calculated in DIVA. The first one is the Useful Daylight Illuminance (UDI). UDI is defined as [16]:

> the annual occurrence of illuminances across the work plane that are within a range considered "useful" by occupants. The range considered "useful" is based on a survey of reports of occupant preferences and behaviour in daylit offices with user operated shading devices. Daylight illuminances in the range 100-300 lux are considered effective either as the sole source of illumination or in conjunction with artificial lighting. Daylight illuminances in the range 300 to around 3,000 lux are often perceived either as desirable or at least tolerable.

We compute UDI values through a time-consuming simulation. The output is represented as false-color map.

The second metric is glare. An example of the glare values over one year, with time slots of one hour, is provided in Figure 2. The glare values are in the range $[0,1]$, and values greater than or equal to 0.45 are considered intolerable.

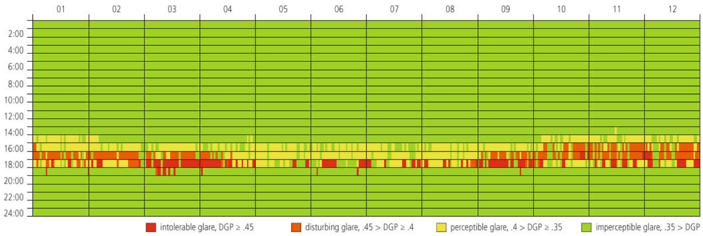


Fig. 2 Example of the output of a simulation for glare on a room over one year

We define our objective function in terms of UDI and glare. Let s be the output of a lighting simulation. Let $U(s)$ be the value of UDI normalized to be in the range $[0,1]$, $G(s)$ be a value in $[0,1]$ expressing the average quality of the glare values over one year, and $T(s)$ equal to 1 if there are too many large (intolerable) glare values, and 0 otherwise. The objective function we want to maximize is expressed as:

$$f(s) = \alpha U(s) + (1-\alpha)G(s) - \alpha(1-\alpha)|U(s) - G(s)| - T(s), \quad \alpha \in (0,1), \quad (1)$$

where the term $\alpha(1-\alpha)|U(s) - G(s)|$ is used to penalize solutions presenting good values of UDI and bad values of glare, or vice-versa. The parameter α is used to decide the relative weights of glare and UDI for the considered room, and this decision is taken by the architects. The values of $f(s)$ are in the range $[-1,1]$, and negative solutions are assumed infeasible. Notice that $f(s)$ is not defined in terms of the design parameters (i.e., the angles of the disk-shaped louvers), but it can be computed from the output of the simulations.

Our goal is to optimize the design parameters (opening angles of the disk-shaped louvers) to achieve an optimal value for this objective function. Note that a designer can change the relative importance of brightness versus glare, for example because the function of the room has changed, and build a new model of the objective function and hot-start the optimization process without re-running the simulations. The designer can thus explore different design scenarios with a minimal need for time-intensive computations. This flexibility is key in the architectural design process, which typically involves not only changes to the solution but also changes to the design objectives [20].

3 The Radial Basis Function Method

The problem discussed above can formally be cast into the following general form:

$$\min f(x), \text{ subject to } x \in [x^L, x^U], \quad (2)$$

where $f : \mathbb{R}^n \to \mathbb{R}$, and $x^L, x^U \in \mathbb{R}^n$ are vectors of lower and upper bounds on the decision variables that define the design space. The analytical expression for f is unknown and function values are only available through an oracle. We assume that

f is continuous with respect to all variables. The method we discuss can be extended to take into account integer variables and additional (explicit) constraints to enforce known relationships between the decision variables, but we keep this simplified formulation for ease of exposition. Our implementation, however, can handle integer variables.

The main idea of the RBF method proposed in [6] is to use radial basis functions (i.e., real-valued functions whose value only depends on the distance from a center x_i, that is, functions of the form $\phi(\|x - x_i\|)$) to build a surrogate model that interpolates the known points. The next evaluation point is chosen by selecting a target objective function value f_k^*, and finding the point in the design space that minimizes the "bumpiness" of the resulting interpolant, if we were to add an interpolation node at the chosen target level. Thanks to the properties of RBFs, we can find an analytical expression for a measure of bumpiness. For space reasons, here we provide a brief overview of the method, skipping most of the details. The interested reader can refer to [6, 9, 19].

Let $\Omega := \{x \in [x^L, x^U]\} \subset \mathbb{R}^n$. Given k distinct points $x_1, \ldots, x_k \in \Omega$, the RBF interpolant s_k is defined as:

$$s_k(x) := \sum_{i=1}^{k} \lambda_i (\|x - x_i\|)^3 + ax + b, \tag{3}$$

where $\lambda_1, \ldots, \lambda_k \in \mathbb{R}$ are the coefficients of the RBFs, $\|\cdot\|$ is the Euclidean norm, $a \in \mathbb{R}^n, b \in \mathbb{R}$. In the expression above we are using the RBF $\phi : \mathbb{R}_+ \rightarrow \mathbb{R}, \phi(r) = r^3$. In this paper we only discuss the cubic RBF $\phi(r) = r^3$, which gives the best performance in our computational experiments, but in general other choices of the RBF are possible (e.g., thin-plate splines). The parameters λ, a, b of the function $s_k(x)$ that interpolates the k points $(x_i, f(x_i)), \forall i \in \{1, \ldots, k\}$ can be determined by solving a linear system (see e.g., [6]). In the following, we will keep referring to s_k as our surrogate model of f after k function evaluations. An overview of the algorithm is given in Figure 3.

This methods finds a solution within ε of the global optimum, provided that the unknown function is sufficiently smooth, the algorithm is executed for a large enough number of iterations, and the choice of target values f_k^* satisfies a technical condition [6].

We still need to address the following questions:

- How do we choose the initial points needed to create the first interpolant?
- How do we choose the next point to evaluate through the oracle using a target function value?

3.1 Selection of the Initial Points

The initial sample points must be linearly independent to guarantee existence of an interpolant. A commonly used strategy [6, 8, 9] is to choose corner points of the box Ω. Taking all the 2^n corner points is impractical if n is large (remember that

```
input : oracle for $f$, domain $[x^L, x^U]$, maximum # evaluations $n_{\max}$
output: best solution found within $n_{\max}$ evaluations

evaluate $f$ at $k_0$ starting points $x_1 \dots x_{k0}$;
$i \longleftarrow \arg\min\{f(x_i), \forall i \in \{1, \dots, k_0\}\}$;
$(x^*, f^*) \longleftarrow (x_i, f(x_i))$;
$k \longleftarrow k_0$;
while $k < n_{\max}$ do
    compute the interpolant $s_k(x)$ using the $k$ points evaluated so far;
    choose a target value $f_k^*$ and select the next evaluation point $x_{k+1}$;
    evaluate $f(x_{k+1})$ through the oracle;
    if $f(x_{k+1}) < f^*$ then
        $x^* \longleftarrow x_{k+1}$;
        $f^* \longleftarrow f(x_{k+1})$;
    end
    $k \longleftarrow k+1$;
end
return $(x^*, f^*)$
```

Fig. 3 The RBF method

oracle is evaluated at each of these points). A commonly used strategy is to use a Latin Hypercube experimental design of size $n+1$; this is our default strategy in practice. Other choices are possible, see e.g. [9]. Points sampled according to these strategies may not be feasible if there are integer variables and/or additional explicit constraints. [9] suggests sampling more points than strictly necessary (i.e., $> n+1$), and picking the first $n+1$ feasible ones, rounding integer variables if necessary. In practice, feasibility in this case must not be too difficult to obtain, otherwise solving the initial problem (2) is essentially hopeless.

3.2 Selection of the Next Evaluation Point

Arguably, the most important aspect of the algorithm is how to choose the next evaluation point. The idea is to define a target objective function value f_k^* lower than the best known objective value f^*, and find the point x_{k+1} in the domain where the surrogate model is more likely to take the value f_k^*. More precisely, this point will be the one for which the interpolant $s_{k+1}(x)$ that interpolates the previous points $(x_i, f(x_i))$, $\forall i \in \{1, \dots, k\}$ and (x_{k+1}, f_k^*), is the least “bumpy”.

We give an example in Figure 4. Given some points at which the function value is known (blue circles), i.e., the previous iterates, and a target objective function value f_k^* (the dashed line), we can obtain different surrogate models depending on the choice of the point x_{k+1} where the function should assume value f_k^*. In the example, the green surrogate model is less bumpy than the red one. The next evaluation point will be the one yielding the least bumpy surrogate model reaching the value f_k^*. Intuitively, this favors the “simplest” model that interpolates the current points and achieves the target f_k^*.

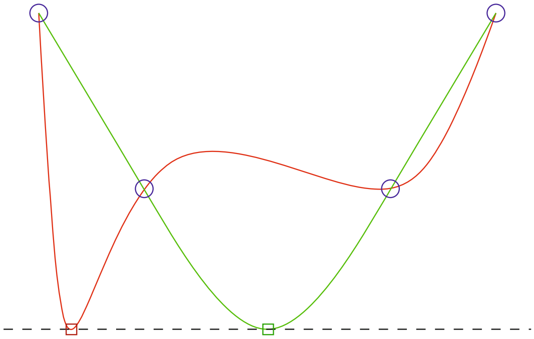


Fig. 4 Example showing two interpolants obtained by choosing a different value of x_{k+1} for a given target value f_k^* (dashed line). The red surrogate model is more bumpy than the green one.

As stated earlier, we can define a measure of bumpiness using the properties of RBFs. Indeed, we can write a nonlinear, nonconvex optimization problem to compute the point that yields the least bumpy interpolant at function value f_k^* [6, 9]. Even though solving this auxiliary problem is not easy in general, we have access to first and second order information, hence it is not as difficult as the original black-box problem. Moreover, it is usually sufficient to find a good local minimum.

We still need to specify how to choose the target objective function value f_k^*. In order to guarantee the convergence of the method, [6] proposes a cyclic strategy to choose f_k^*:

- **inf step**: the target objective function value is set to $-\infty$. The new point x_{k+1} minimizing the bumpiness will be in a large gap between existing interpolation points. This is an *exploration* phase: the algorithm tries to improve the surrogate model in unknown parts of the domain.
- **local step**: the target objective function value is set to $\min_x s_k(x)$, hence the new evaluation point x_{k+1} will be the minimum of the interpolant. This is an *exploitation* phase: we try to find the best objective function value based on the current surrogate model.
- **global step**: the target objective function value is between the two extremes $-\infty$ and $\min_x s_k(x)$. In this case, there is a balance between improving the model quality and finding the minimum.

3.3 *Extensions*

We describe here some extensions to the basic RBF algorithm. Our preliminary computational evaluations shows that these can significantly improve performance of the method, but for space reasons we only give a brief overview.

3.3.1 A Measure of Model Quality

One of the drawbacks of the RBF method is that there is no mechanism to assess the accuracy of the surrogate model. The algorithm uses a cyclic search strategy that oscillates between global search and local search (possibly with intermediate steps), regardless of model quality. However, one can expect that if the current model is a poor approximation of the true objective function, the local search phase is unlikely to find a point with a good objective function value.

We propose to assess model quality using a cross validation scheme. Cross validation is a commonly used model validation technique in statistics and machine learning, and is discussed virtually by any textbook in the field. Given a data set, cross validation consists in using part of the data set to fit a model, and testing its quality on the remaining data. The process is then iterated by choosing different parts of the data set for model fitting and for testing.

We say that a measure of model quality $\mathscr{M}$ is *convergent* if $\lim_{k\to\infty} \mathscr{M}(s_k) = \lim_{k\to\infty} \int |s_k(x) - f(x)| \,\mathrm{d}x$. This implies that $\mathscr{M}$ converges to a reliable measure of difference between the surrogate model and the true objective function f. We are implicitly assuming that the integral exists and is well-defined. It is not hard to prove that $\lim_{k\to\infty} \int |s_k(x) - f(x)| \,\mathrm{d}x = 0$, due to properties of the algorithm. It is easy to design convergent measures, but from a practical point of view we are interested in how closely the sequence $\{\mathscr{M}(s_k)\}_k$ follows $\{\int |s_k(x) - f(x)| \,\mathrm{d}x\}_k$.

The method we advocate is as follows. Remember that function values are known at k points $x_1,\dots,x_k$. Then, for $j \in \{1,\dots,k\}$ we can fit a surrogate model $\tilde{s}_k$ to the points $\{x_i, \forall i \in \{1,\dots,k\} \,|\, i \neq j\}$, and evaluate the peformance at $(x_j, f(x_j))$. In particular, we compute the value $|\tilde{s}_k(x_j) - f(x_j)|$, which is zero if the model is a perfect fit, and > 0 otherwise. We can then average these statistics for $j \in \{1,\dots,k\}$ to compute an overall model quality score: this will be our measure $\mathscr{M}$. Our approach can be categorized as leave-one-out cross-validation. While leave-one-out cross-validation is often computationally expensive, we can show that $\mathscr{M}$ can be computed efficiently by solving a sequence of k linear programs, where each linear program in the sequence can be warmstarted.

We use this measure of quality to dynamically select among different surrogate models of the objective functions, obtained using different RBFs. In particular, periodically during the optimization process we select which model seems to yield the best predictive power globally (for the global search phase of the algorithm) and locally around the optimum (for the local search phase of the algorithm).

Furthermore, at the end of the algorithm we can output the surrogate model and a corresponding “confidence level” to quantify how reliable we expect the surrogate model to be. This has important practical implications: the surrogate model can be used to perform sensitivity analysis around the optimal solution, i.e., explore how the objective function changes if the design parameters are perturbed, without having to resort to new (expensive) oracle evaluations; see the discussion in Section 2. The confidence level informs the user about the reliability of such sensitivity analysis.

3.3.2 Noisy Functions Evaluation

In the application context studied in this paper, there is a trade-off between computing time and the accuracy of the black-box function f. This is a common situation

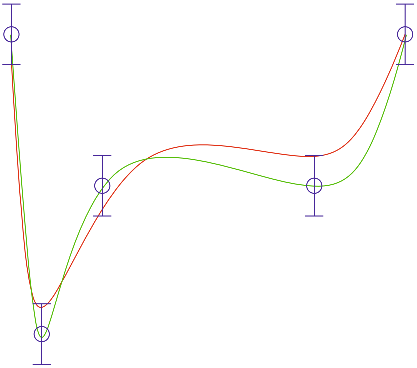


Fig. 5 Function evaluations affected by errors: the values returned by LQ simulations are the blue dots. The green function interpolates exactly at those points, the red function is a less bumpy interpolant, and is still within the allowed error tolerances.

in practice: the simulation software used to compute f can often be parameterized to achieve different levels of accuracy. Our target application has this characteristic.

To speed-up the optimization process, we would like to exploit low accuracy and therefore faster simulations. In the following, we will denote by LQ ("low quality") the low accuracy, high speed setting, and by HQ ("high quality") the high accuracy, low speed setting. Our assumption is that function evaluations computed in LQ are affected by an unknown error, whereas function evaluations computed in HQ represent the true objective function $f(x)$. The specifics of what LQ and HQ correspond to (and as a consequence the estimation of the error) are application-dependent. Further, we assume that we have an estimate of the maximum relative (or absolute) error ε_r. Such an estimate is typically based on domain knowledge. Notice that the surrogate model interpolating LQ points should be allowed to deviate from the values returned by the oracle by an amount within ε_r. Consider the example of Figure 5: if the points were HQ, i.e., considered "exact", the interpolant would be the green function. If we assume that function evaluations are affected by noise, and the true function values may be located anywhere in an interval around their true value, then we prefer the red surrogate model, because it is the least bumpy among the models that interpolate within the given noise intervals. We can show that the RBF surrogate model under noisy function evaluations can be determined by solving a convex quadratic problem. We advocate the following approach: at a first stage we solve (2) in LQ until a termination condition is met, based on number of function evaluation, an estimation of model quality (for example employing the cross validation explained in the previous section), or stalling. At a second stage (2) is reoptimized performing additional function evaluations in HQ, still allowing the LQ function evaluations to vary by at most ε_r. During the local search phase, if the minimum of the interpolant occurs at a LQ point, we re-evaluate the same point in HQ.

3.4 Open-Source Implementation

We implemented the RBF method for MATLAB$^{\text{TM}}$/Octave and Python. The former requires AMPL® license [5]. Our implementation is open-source and is available

Table 1 Number of function evaluations needed to find a value within 1% from the global optimum for different black-box methods on instances from the Dixon-Szegö test set

Function	Dimension	Our RBF	TOMLAB RBF	DE
Branin	2	36	26	1190
Goldstein-Price	2	30	27	1018
Hartman 3	3	42	22	476
Hartman 6	6	114	87	7220
Shekel 5	4	116	96	6400

from the authors[1]. To solve the nonlinear (mixed-integer in the presence of integer variables) optimization problems generated during the various steps of the algorithm a nonlinear solver is also needed. In our tests we use BONMIN [2] with the nonlinear solver IPOPT [22]. Currently, our algorithm can handle problems in any dimension, and it can handle both continuous and integer variables. Clearly an increase of the problem dimension significantly increases the difficulty of finding a global optimum. Besides our own version, we are aware of only one available implementation of the RBF method: the commercial TOMLAB® toolkit for MATLAB™.

To give an idea of the performance of our implementation, we show in Table 1 the number of iterations needed to find a solution within 1% from the global optimum for some well-known instances of the literature (i.e., the Dixon-Szegö [4]) having multiple global and local optima and dimension up to 6. We compare the number of evaluations of the oracle of our method, the best results obtained by TOMLAB® as reported by [8], and an evolutionary algorithm (DE [21]). We remark that results for these test were obtained with the default parameters of our implementation, i.e., an "out-of-the-box" run without any sort of tuning. The number of function evaluations is comparable to the best results of TOMLAB® as given by [8]. It is more than one order of magnitude better than an evolutionary algorithm, which is common practice for optimization in architectural design.

To summarize, our open-source implementation is competitive with state-of-the-art commercial software, even without any parameter tuning. At the present stage, the integration with architectural design software is still in progress. However, we believe that this paper shows that the application of state-of-the-art optimization methods in the architectural design context could prove an invaluable decision-making tool.

References

1. Björkman, M., Holmström, K.: Global optimization of costly nonconvex functions using radial basis functions. Optimization and Engineering 1(4), 373–397 (2000)
2. Bonami, P., Lee, J.: BONMIN user's manual. Tech. rep. IBM Corporation (2007)

---

[1] At the moment of writing this paper, the license terms are still to be defined, but we expect a BSD-derivate that will be free for academic use.

3. Delanda, M.: Deleuze and the use of the genetic algorithm in architecture. In: Contemporary Techniques in Architecture, pp. 9–11. Wiley-Academy, London (2002)
4. Dixon, L., Szego, G.: The global optimization problem: an introduction. In: Dixon, L., Szego, G. (eds.) Towards Global Optimization, pp. 1–15. North Holland, Amsterdam (1975)
5. Fourer, R., Gay, D.: The AMPL Book. Duxbury Press, Pacific Grove (2002)
6. Gutmann, H.M.: A radial basis function method for global optimization. Journal of Global Optimization 19, 201–227 (2001)
7. Hemker, T.: Derivative free surrogate optimization for mixed-integer nonlinear blackbox problems in engineering. Master's thesis, Technischen Universität Darmstadt (2008)
8. Holmström, K.: An adaptive radial basis algorithm (ARBF) for expensive black-box global optimization. Journal of Global Optimization 41(3), 447–464 (2008)
9. Holmström, K., Quttineh, N.H., Edvall, M.M.: An adaptive radial basis algorithm (ARBF) for expensive black-box mixed-integer constrained global optimization. Optimization and Engineering 9(4), 311–339 (2008)
10. Jakubiec, J.A., Reinhart, C.F.: DIVA 2.0: Integrating daylight and thermal simulations using Rhinoceros 3d, DAYSIM and EnergyPlus. In: Proceedings of the 12th International IBPSA Conference on Building Simulation, pp. 2202–2209 (2011)
11. Jones, D., Schonlau, M., Welch, W.: Efficient global optimization of expensive black-box functions. Journal of Global Optimization 13(4), 455–492 (1998)
12. Kaveh, A., Laknejadi, K.: A new multi-swarm multi-objective optimization method for structural design. Advances in Engineering Software 58, 54–69 (2013)
13. Kicinger, R., Arciszewski, T., Jong, K.D.: Evolutionary computation and structural design: A survey of the state-of-the-art. Computers and Structures 83(23-24), 1943–1978 (2005)
14. Krob, D.: Modelling of complex software systems: A reasoned overview. In: Najm, E., Pradat-Peyre, J.-F., Donzeau-Gouge, V.V. (eds.) FORTE 2006. LNCS, vol. 4229, pp. 1–22. Springer, Heidelberg (2006)
15. Lin, S.H.E., Gerber, D.J.: Designing-in performance: A framework for evolutionary energy performance feedback in early stage design. Automation in Construction 38, 59–73 (2014)
16. Mardaljevic, J., Andersen, M., Roy, N., Christoffersen, J.: Daylighting, Artificial Lighting and Non-Visual Effects Study for a Residential Building. Tech. rep., Loughborough, UK (2012)
17. Miles, J.: Genetic algorithms for design. In: Waszczyszyn, Z. (ed.) Advances of Soft Computing in Engineering, pp. 1–56. Springer, Vienna (2010)
18. Oxman, R.: Performance-based design: current practices and research issues. International Journal of Architectural Computing 6(1), 1–17 (2008)
19. Regis, R., Shoemaker, C.: Improved strategies for radial basis function methods for global optimization. Journal of Global Optimization 37, 113–135 (2007)
20. Rittel, H.W.J., Webber, M.M.: Dilemmas in a general theory of planning. Policy Sciences 4(2), 155–169 (1973)
21. Storn, R., Price, K.: Differential evolution – a simple and efficient heuristic for global optimization over continuous spaces. Journal of Global Optimization 11, 341–359 (1997)
22. Wächter, A., Biegler, L.T.: On the implementation of a primal-dual interior point filter line search algorithm for large-scale nonlinear programming. Mathematical Programming 106(1), 25–57 (2006)
23. Woodbury, R.F., Burrow, A.L.: Whither design space? AI EDAM 20(2), 63–82 (2006)

Aligning Cyber-Physical System Safety and Security

Giedre Sabaliauskaite and Aditya P. Mathur

Abstract. Safety and security are two key properties of Cyber-Physical Systems (CPS). Safety is aimed at protecting the systems from accidental failures in order to avoid hazards, while security is focused on protecting the systems from intentional attacks. They share identical goals – protecting CPS from failing. When aligned within a CPS, safety and security work well together in providing a solid foundation of an invincible CPS, while weak alignment may produce inefficient development and partially-protected systems. The need of such alignment has been recognized by the research community, the industry, as well as the International Society of Automation (ISA), which identified a need of alignment between safety and security standards ISA84 (IEC 61511) and ISA99 (IEC 62443). We propose an approach for aligning CPS safety and security at early development phases by synchronizing safety and security lifecycles based on ISA84 and ISA99 standards. The alignment is achieved by merging safety and security lifecycle phases, and developing an unified model – Failure-Attack-CounTermeasure (FACT) Graph. The FACT graph incorporates safety artefacts (fault trees and safety countermeasures) and security artefacts (attack trees and security countermeasures), and can be used during safety and security alignment analysis, as well as in later CPS development and operation phases, such as verification, validation, monitoring, and periodic safety and security assessment.

Keywords: Cyber-physical systems, CPS, Safety, Security, Alignment, ISA84; IEC 61511, ISA99, IEC 62443, Fault trees, Attack trees.

Giedre Sabaliauskaite · Aditya P. Mathur
Information Systems Technology & Design Pillar,
Singapore University of Technology and Design,
Singapore
e-mail: `{giedre,aditya_mathur}@sutd.edu.sg`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_4

1 Introduction

Safety and security are two key properties of Cyber-Physical Systems (CPS) [6, 15]. They share identical goals – protecting CPS from failures [17]. Safety is aimed at protecting the systems from accidental failures in order to avoid hazards, while security is focusing on protecting the systems from intentional attacks. Safety and security are particularly important in industrial control systems, where hazards include explosions, fires, floods, chemical/biochemical spills and releases, potential crashes of vehicles, etc.

In order to protect CPS from failing, safety and security have to be aligned and their 'activities or systems organised so that they match or fit well together' (MacMillan Dictionary's definition of 'align'). When aligned within a cyber-physical system, safety and security work well together in providing a solid foundation of an invincible CPS.

Weak alignment between security and safety may produce inefficient development and partially-protected systems. For example, excess costs could be spent on redundant safety and security countermeasures. Furthermore, security countermeasures may weaken CPS safety, or vice versa – safety countermeasures may weaken security [14]. If there is no alignment between safety and security countermeasures, these interdependencies are not detected in the early system development phases and may lead to a number of problems that affect later CPS development or even operation phases.

Over the years, separate research communities have dealt with threats to security versus safety [16]. Two international standards have been proposed by the International Society of Automation (ISA) to address CPS safety and security needs: ISA84 standard (also called IEC 61511) on safety instrumented systems [5], and ISA99 standard (also called IEC 62443) on control system security [8].

As systems are becoming more complex and integrated, the distinction between safety and security is beginning to weaken. Researchers are starting to recognize a need of collaboration between these two communities [2, 16]. ISA has also identified a need of alignment between safety and security, and formed a working group, Work Group 7 - Safety and Security, to investigate alignment and common issues between security and safety [10].

Some techniques and approaches from the safety domain have already been adapted for security and vice versa. The next step is development of techniques and approaches for integrated improvement of both safety and security [16].

Identifying and characterizing safety and security interdependencies in the early stages of CPS development (specification and design) in order to manage their consequences and optimize operational resources and system performance is still a challenge, which needs to be addressed [14]. Furthermore, the proposed solutions need to be compliant with industrial standards for CPS safety and security.

In this paper, we propose an approach for aligning CPS safety and security lifecycles, based on ISA84 and ISA99 standards. The alignment is achieved by merging safety and security lifecycle phases and developing an unified model, Failure-Attack-CounTermeasure (FACT) Graph. The FACT graph incorporates safety artefacts (fault trees and safety countermeasures) and security artefacts (attack trees and security countermeasures), and can be used for safety and security alignment analysis.

The proposed approach and the FACT graph not only help to assure that CPS safety (ISA84) and security (ISA99) are implemented in a consistent way, but also enable organizations analyze system vulnerabilities to newly identified accidental and intentional failures during CPS operation, and update countermeasure set in order to provide required level of safety and security.

The remainder of the paper is organized as follows. Section 2 describes the related work. An integrated safety and security lifecycle process is presented in Section 3. Section 4 describes a CPS safety and security alignment model – FACT graph. An example of the FACT graph is presented in Section 5. Finally, Section 6 concludes the paper.

2 Related Work

2.1 CPS Safety

Safety concerns of the process industries can be addressed by the use of the ISA84 standard, which describes the application of Safety Instrumented Systems (SIS) to achieve and/or to maintain a safe state of the process [5]. A SIS is aimed at performing specific control functions to maintain safe operation of the process when predefined conditions are violated [5].

Safety life-cycle consists of the following phases (see Fig. 4, left side): Hazard and risk assessment; Allocation of safety functions to protection layers; Safety requirements specification; Design of safety countermeasures; Installation, commissioning and validation; Operation and maintenance; Modification; and Decommissioning [5].

The main objectives of the hazard and risk assessment phase are to determine the hazards and hazardous events of the process and associated equipment, the process risks associated with the hazardous events, and the safety functions to achieve the necessary risk reduction.

In the second phase, safety functions are assigned to protection layers. Then, the required safety instrumented systems and associated safety integrity levels are determined. In the safety requirement specification phase, SIS specific safety requirements are derived from the overall CPS safety requirements, defined during the hazard and risk assessment phase.

In the next phase, safety countermeasure – SIS and other means of risk reduction – are designed taking into account the safety requirements. Then, security countermeasures are implemented, validated, and maintained during the CPS operation.

Typical artefacts, developed/defined throughout safety life-cycle, are the following among others (see Fig. 4): failure initiating events, fault trees, safety functions, protection layers, safety requirements and safety countermeasures.

Fault Tree Analysis (FTA) [7] is a technique widely used for hazard and risk assessment. The purpose of the technique is to graphically present the possible normal and faulty events that can cause the top-level undesired event. The fault tree consists of the following components: nodes (undesired events in the system), gates (relations between nodes; can be AND or OR gates), and edges (path of the undesired events through the system).

For the reason that FTA is a graphical technique widely used by researchers and practitioners, we will employ it in our safety-security alignment approach.

In order to illustrate fault tree analysis, we borrowed an example of a pressurized vessel control system from [5]. Pressurized vessel control process is shown in Fig. 1, and its fault tree is presented in Fig. 2.

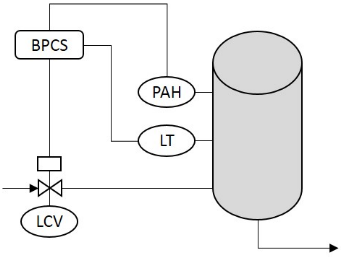


Fig. 1 Pressurized vessel process [5]

The process comprises of a pressurized vessel containing volatile flammable liquid [5]. Basic Process Control System (BPCS) is controlling the process by monitoring the signal from the Level Transmitter (LT) and controlling the operation of the Level Control Valve (LCV) (see Fig. 1). The system is equipped with a high pressure alarm – Pressure Alarm High (PAH), which is initiated if level transmitter measures high pressure level in order to alert the operator to take appropriate action to stop inflow of material.

A fault tree shown in Fig. 2 identifies the events, which contribute to the development of overpressure condition in the vessel [5]. The top event, Overpressurization, is caused either by failure of the basic process control system function, or an external event, such as e.g. fire. BPCS function failure can be caused either by BPCS failure, or field device failures, such as sensor or actuator (valve) failure. Two transfer gates, included into the tree, indicate connections to related fault trees – external event fault tree and BPCS failure fault tree (these trees are not included in this paper).

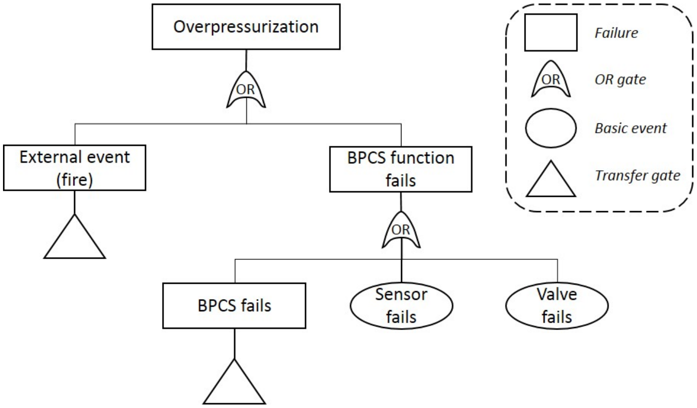


Fig. 2 Fault tree of overpressure of the vessel [5]

2.2 CPS Security

To address cyber-physical system security needs, ISA developed an ISA99 standard "Security for Industrial Automation and Control Systems" [8]. The primary goal of this standard is to provide a flexible framework that facilitates addressing CPS vulnerabilities and applying necessary countermeasures in systematic, defensible manner.

ISA99 describes a generic view of an integrated manufacturing or production CPS, expressed as a series of logic levels [8]: Enterprise Business Systems, Operations Management, Supervisory Control, Basic Control, and Physical Process.

In large or complex system it may not be practical or necessary to apply the same security level to all components. ISA99 proposes to divide systems in zones and conduits in order to meet the security goals [8]. Security zone is a logical or physical grouping of physical, informational, and application assets sharing common security requirements. Zones can be formed at different system logical levels, or across levels. Conduits are used to define communications, and may connect entities within zones, or may connect different zones.

Security lifecycle consists of the three main phases: Assessment, Implementation and Maintenance [8]. The assessment phase consists of Process risk assessment, Security requirements specification, Zone and conduit identification, and Risk assessment for each zone and conduit (see Fig. 4, right side). During the implementation phase, security countermeasures are designed, validated, developed and verified. Finally, maintenance phase includes operation and maintenance, security monitoring and periodic assessment, and modification and decommissioning.

Typical artefacts, developed/defined throughout security lifecycle process, include the following items among others: attack trees, security requirements, security zones and conduits, and security countermeasures.

Integrity, availability and confidentiality are three high-level cyber security objectives for CPS [9]. A lack of confidentiality results in disclosure, when an unauthorized entity gains access to data. A lack of integrity leads to deception – when an authorized party receives false data and believes it is true. While a lack of availability results in denial of service (DoS) when an authorized entity cannot receive commands or data. Deception, disclosure and DoS are three basic types of cyber-attacks on CPS [9].

Attack trees [3, 11] are widely used for security risk assessment. Attack tree is a graph that describes the steps of attack process. It uses the same basic symbols as fault trees: nodes (represent attacks), gates (AND and OR gates), and edges (path of attacks through the system). Several authors propose to use additional symbols in attack trees. E.g. dynamic, "trigger" edges [12] can be used in situations when one attack event (e.g. Attack 1) triggers the other (e.g. Attack 2). In this case, Attack 2 can be realizable only if Attack 1 has been completed.

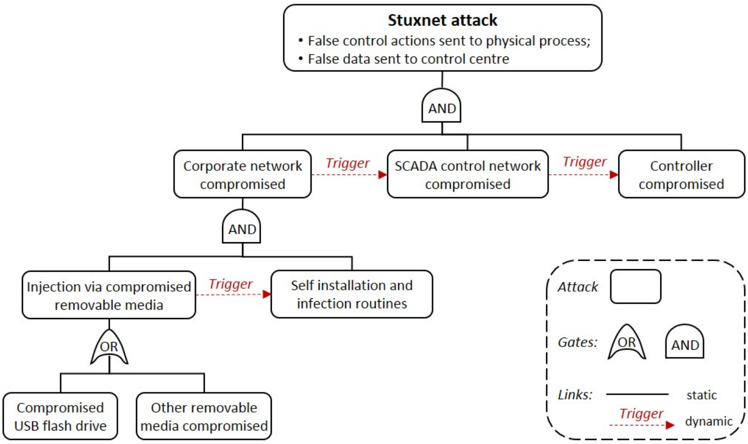


Fig. 3 Attack tree example – Stuxnet

An example of an attack tree is shown in Fig. 3. It depicts attack process steps of the Stuxnet attack [1]. The goal of the Stuxnet attack is to compromise controller, which is controlling a SCADA system. The attack starts with injection via compromised removable media, which could be done by user opening a compromised file folder either on USB flash drive, or other removable media. Once this step is completed, Stuxnet worm instantaneously starts its self-installation and infection routines, thus there is a "trigger" line between these nodes. After injection and self-installation and infection routines are completed, an attacker is able to compromise corporate network, which allows him to gain access to SCADA control network, and eventually to compromise a controller.

Several authors proposed to add defense mechanisms to the attack trees [3, 11]. Defense nodes (security countermeasures) can be attached either to attack leafs [11], or to any node in an attack tree, as defined in the Attack Countermeasure

Tree (ACT) approach [3]. In an ACT approach, once attack tree is constructed, and possible security countermeasures are attached to the attack nodes, security analysts can select a set of security countermeasures for implementation, considering a given budget [3].

2.3 *Safety and Security Integration*

Safety and security are interdependent, and these dependencies have to be considered during CPS design phase. There are four types of interdependencies between safety and security [14]: 1) conditional dependencies – security is a condition for safety and vice versa; 2) reinforcement – safety and security countermeasures can strengthen each other; 3) antagonism – they can weaken each other; and 4) independence – no interaction between safety and security.

Several techniques have been proposed in literature for integrating safety and security. They include unified risk definition [2], safety and security life cycle model [17], detecting conflicts between safety and security requirements [13], and integrated graphical model [4] among others.

In [2], authors recommend to expand definition of the safety term hazard to include security related risks. A new definition of mishap is proposed. A mishap is an unplanned event, or series of events, that result in death, injury, occupational illness, or other harm to individual's well-being; damage to or loss of equipment or property; or harm to an organization. These events include system, equipment or component failures, design flaws, user errors, intentional attacks, etc.

A life cycle model for integrated safety and security in automation systems is described in [17]. It focusses on resolving conflicts between safety and security at the requirements and functional levels. Such a life cycle model and its conflict resolution framework are the basis of combining formerly separated networks for safety (e.g. fire alarm system), security (e.g. access control) and operation.

In [13], authors propose a technique for detecting conflicts between safety and security requirements, which helps alleviate contradictory requirements. A tool is used to detect truly coupled requirements among two domains (safety and security), providing system designers with information on requirement contradictions.

In [4], a method for quantitative security risk assessment of complex systems is proposed. It combines fault trees with attack trees by integrating attack trees into pre-existent fault trees. The proposed approach allows considering the interaction of malicious attacks with random failures.

In 2009, ISA formed a joint working group to promote awareness of the impact of cyber-security issues on the safe operation of industrial processes – ISA99 Work Group 7 [10]. To the best of our knowledge, there are no work results published by this group yet.

Despite the existing solutions, there is still a need of an approach, which would help to design safe and secure CPS [14], and which would be compliant with industrial standards for CPS safety and security.

3 Integrated CPS Safety and Security Lifecycle Process

We propose a CPS safety and security alignment approach. The alignment is achieved by merging safety and security lifecycle phases and developing a unified model, failure-attack-countermeasure (FACT) graph (see Fig. 4).

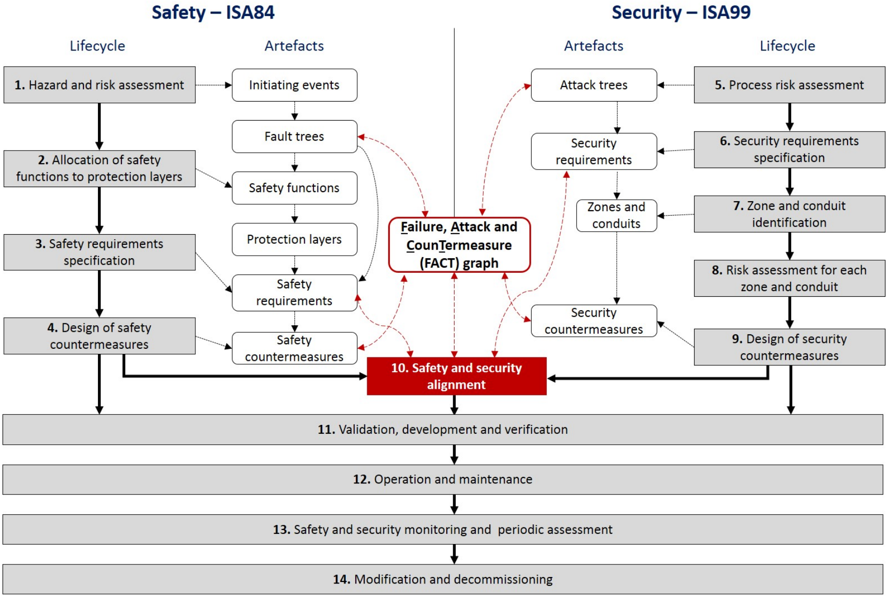


Fig. 4 Merged ISA84 and ISA99 lifecycles

The merged safety and security lifecycle model, which consists of 14 phases, is shown in Fig. 4. The process starts with CPS safety risk assessment and design phases (phases 1 – 4), borrowed from ISA84, followed by security risk assessment and design phases (phases 5 – 9), taken from ISA99. In phase 10, the alignment between safety and security is performed. Finally, phases 11-14 are the merged phases of ISA84 and ISA99 lifecycles and include validation, development, and verification, operation and maintenance, safety and security monitoring and periodic assessment, and modification and decommissioning related activities.

The failure-attack-countermeasure (FACT) graph is formed throughout phases 1-9 of the merged safety and security lifecycle. It incorporates various artefacts: safety artefacts (fault trees and safety countermeasures) and security artefacts (attack trees and security countermeasures). This graph is useful not only for evaluating safety and security alignment in phase 10, but also for safety and security verification and validation in phase 11, and for monitoring and periodic assessment in phase 13. The alignment phase (phase 10) and the FACT graph are described in more detail in Section 4.

4 CPS Safety and Security Alignment Using Failure Attack-Countermeasure Graph

As it has been mentioned in Section 3, FACT graph is formed throughout phases 1-9 of the merged safety and security lifecycle process. Various artefacts from these phases are used as inputs for the FACT graph as shown in Fig. 5: fault trees, safety countermeasures, attack trees, and security countermeasures. Furthermore, security requirements and safety requirements along with FACT graph are the inputs for safety and security alignment analysis.

The FACT graph construction consists of the following four steps:

Step 1. The construction of the graph starts with importing failure trees at the end of the safety hazard and risk assessment phase (phase 1). Whenever possible, interrelated fault trees are connected, using AND or OR gates, in order to provide a complete view of possible failures of the system. These fault trees form a frame of the FACT graph.

Step 2. As soon as the definition and design of safety countermeasures is completed (phase 4), safety countermeasures are added to the FACT graph. They are attached to the failures they are aimed at preventing. This mapping allows us to see the coverage of safety failures by safety countermeasures.

Step 3. Attack trees, formed during process risk assessment phase (phase 5), are added to the FACT graph. Attacks, related to failures in a FACT graph, are attached to the corresponding safety failures. Attack trees are incorporated into fault trees by the use of OR gate, which indicates that a failure may be caused either by accidental failures, or by intentional attacks.

Step 4. After completion of the security countermeasure design phase (phase 9), security countermeasures are added to the FACT graph. We can use ACT technique [3] in this step, which allows attachment of security countermeasures to any node of the attack tree (see Section 2.2).

The FACT graph, constructed during steps 1-4, is a comprehensive system safety and security model, which shows safety and security artefacts and their relationships. It can be used for safety and security alignment analysis (phase 10). In this phase, it is important for safety analysts to work together with security analysts in order to identify any misalignment, duplicates or missing elements. For this purpose, safety requirements, defined in phase 3, and security requirements, defined in phase 6 of the merged lifecycle, are used (see Fig. 5). In phase 10, FACT graph needs to be reviewed and compared against safety and security requirements to determine if the requirements are satisfied, i.e. if proposed safety and security countermeasures provide the necessary risk reduction to achieve tolerable risks of the CPS.

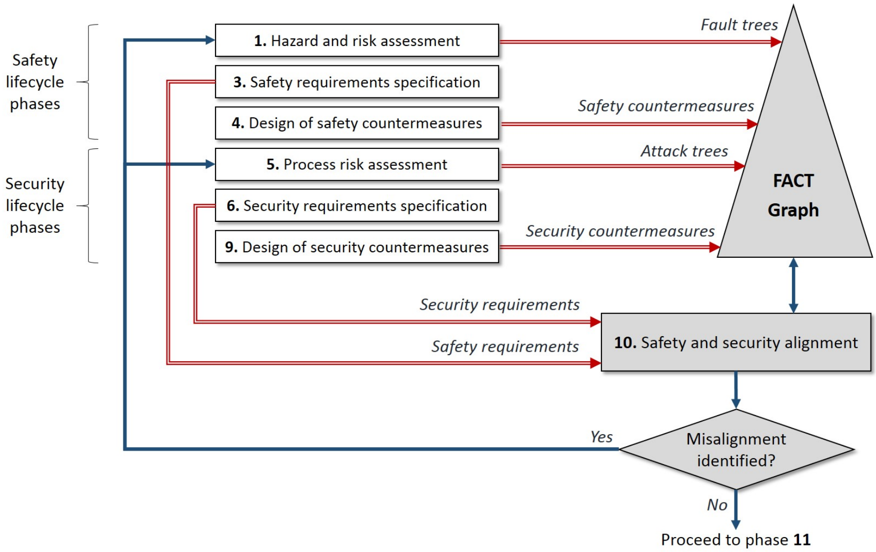


Fig. 5 Safety and security alignment process

Furthermore, FACT graph can be used in later CPS development and maintenance phases as well, e.g. in phase 11 for safety and security verification and validation, and in phase 13 for monitoring and periodic assessment. As new types of failures and attacks are being continuously reported, these newly identified failures and attack should be added to the FACT graph in order to determine if current countermeasure set is sufficient for protecting CPS.

5 FACT Graph Example

This section includes an example of the FACT graph. The FACT graph construction process comprises of four steps, as defined in Section 4. In the first step, fault trees are imported into the FACT graph and their inter-connections are established. In this example, we will use the fault tree of overpressure of the vessel, borrowed from Fig. 2.

In the second step, safety countermeasures are added to the FACT graph. In our example, three safety countermeasures are added, as shown in Fig. 6: SAF1 - a high pressure alarm – Pressure Alarm High (PAH) (see Fig. 1), which is activated if level sensor measures high pressure level in order to alert the operator to take appropriate actions to stop inflow of material; SAF2 – redundant sensor, which could be used in a situation when the primary sensor fails; SAF3 – redundant valve, used in a situation when the primary valve fails.

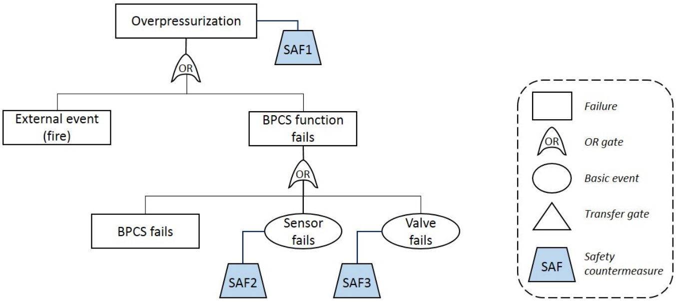


Fig. 6 Failure and safety countermeasure graph of overpressure of the vessel

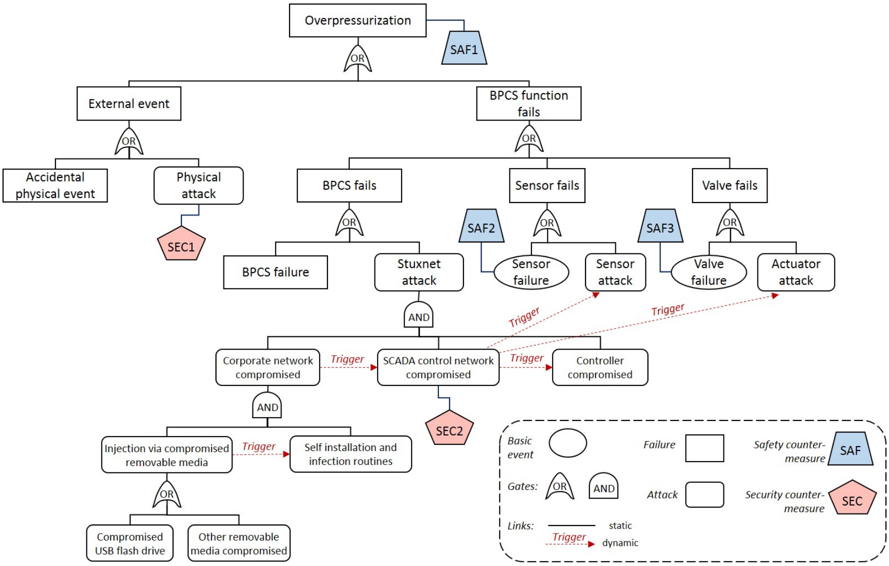


Fig. 7 Failure-Attack-Countermeasure (FACT) graph of overpressure of the vessel

In the third step, attack trees are added to the FACT graph. In our example, Stuxnet attack tree (see Fig. 3) is attached to the failure BPCS Fails using OR gate: BPCS may fail accidentally, or it may be caused by a malicious Stuxnet attack. In addition, three more attacks are added to the graph: Physical attack, Sensor Attack, and Actuator attack. It is important to add inter-connection between attacks to the graph as well. In our example, sensor and actuator attacks may be triggered if SCADA control network is compromised. Thus, links between these attacks are added to the graph (see Fig. 7).

Finally, in the fourth step, security countermeasures are added to the graph. Two security countermeasures are added to the vessel overpressure graph (see Fig. 7): SEC1 – countermeasures against physical attacks to prevent unauthorized access to the equipment, such as security guards, cameras, locks, etc.; SEC2 – countermeasures for detecting SCADA control network attacks. Detection and mitigation of control network attacks is crucial, because it will prevent not only a Stuxnet attack, but also sensor and actuator attacks. Various intrusion detection techniques can be used for detecting attacks on SCADA control networks.

6 Conclusions

In this paper we have proposed an approach for aligning cyber-physical system safety and security at early development phases. The proposed approach suggests a way to integrate safety and security lifecycle process phases, defined by ISA84 and ISA99 standards. Using this approach, practitioners may align CPS safety and security activities, by following the merged 14-phase safety and security lifecycle process (see Fig. 4), during which alignment model – FACT graph – is created.

The FACT graph can be used to identify any misalignment between safety and security countermeasures, as well as countermeasure duplicates and missing means of protection. In FACT graph, safety and security countermeasure are attached to the relevant faults and attacks, thus it is easy to identify interrelated countermeasures and analyze their interdependencies.

FACT graph is built on a frame of interconnected CPS failures and attacks, and it is continuously updated during CPS development and operation phases by adding newly identified failures and attacks to it. This will help to determine if current countermeasures are sufficient to detect and mitigate newly identified failures and attacks. Furthermore, FACT graph will help to identify redundant safety and security countermeasures – countermeasures, which are aimed at preventing the same attack/failure from happening. Safety and security specialist along with managers will be able to select and optimal countermeasure set to provide necessary protection considering a given budget.

Further research is needed evaluate the proposed approach in a real CPS. We are planning to apply the safety and security alignment approach in an industrial process control system, water purification CPS, which we are currently developing.

Acknowledgments. This research is supported by a start-up grant from the Singapore University of Technology and Design, (http://www.sutd.edu.sg). It is a part of ongoing research at the iTrust – Center for Research in Cyber Security.

References

[1] Karnouskos, S.: Stuxnet worm impact on industrial cyber-physical system security. In: Proceedings of the 37th IEEE Annual Conference on Ind. Electronics Soc. (IECON 2011), pp. 4490–4494 (November 2011), doi:10.1109/IECON.2011.6120048

[2] Stoneburner, G.: Toward a Unified Security-Safety Model. Computer 39(8), 96–97 (2006)

[3] Roy, A., Dong, S.K., Trivedi, K.S.: Scalable optimal countermeasure selection using implicit enumeration on attack countermeasure trees. In: Proceedings of the 42nd Annual IEEE/IFIP International Conference on Dependable Systems and Networks (DSN 2012), pp. 1–12 (June 2012)

[4] Nai Fovino, I., Masera, M., De Cian, A.: Integrating cyber attacks within fault trees. Reliability Engineering & System Safety 94(9), 1394–1402 (2009)

[5] ANSI/ISA 84.00.01-2004, Application of Safety Instrumented Systems for the Process Industries. The Instrumentation, Systems, and Automation Society, Research Triangle Park, NC (2004)

[6] Piètre-Cambacédès, L., Chaudet, C.: The SEMA referential framework: Avoiding ambiguities in the terms “security” and “safety”. International Journal of Critical Infrastructures Protection 3(2) (June 2010)

[7] Reichenbach, F., Alme, K.-J., Endresen, J.: On the significance of fault tree analysis in practice. In: Proceedings of the IEEE Conference on Emerging Technologies & Factory Automation (ETFA 2009), September 22-25, pp. 1–7 (2009)

[8] ANSI/ISA-99-00-01-2007. Security for Industrial Automation and Control Systems. Part 1: Terminology, Concepts, and Models. The Instrumentation, Systems, and Automation Society, Research Triangle Park, NC (2007)

[9] Amin, S.: On Cyber Security for Networked Control Systems. PhD Thesis, University of California, Berkeley (2011)

[10] ISA 99 Work Group 7 – Safety and Security (Joint with ISA84 committee), `http://isa99.isa.org/ISA99%20Wiki/WG7.aspx` (references on April 11, 2014)

[11] Ten, C.-W., Liu, C.-C., Govindarasu, M.: Vulnerability Assessment of Cybersecurity for SCADA Systems Using Attack Trees. In: Power Engineering Society General Meeting, pp. 1–8 (June 2007), doi:10.1109/PES.2007.385876

[12] Kriaa, S., Bouissou, M., Pietre-Cambacedes, L.: Modeling the Stuxnet attack with BDMP: Towards more formal risk assessments. In: Proceedings of the 7th International Conference on Risk and Security of Internet and Systems (CRiSIS 2012), pp. 1–8 (October 2012), doi:10.1109/CRISIS.2012.6378942

[13] Sun, M., Mohan, S., Sha, L., Gunter, C.A.: Addressing Safety and Security Contradictions in Cyber-Physical Systems. In: Proceedings of the Workshop on Future Directions in Cyber-Physical Systems (July 2009)

[14] Piètre-Cambacédès, L., Bouissou, M.: Modeling safety and security interdependencies with BDMP (Boolean logic Driven Markov Processes). In: Proceedings of the IEEE International Conference on Systems Man and Cybernetics (SMC 2010), pp. 2852–2861 (October 2010)

[15] Banerjee, A., Venkatasubramanian, K.K., Mukherjee, T., Gupta, S.K.S.: Ensuring Safety, Security, and Sustainability of Mission-Critical Cyber–Physical Systems. Proceedings of the IEEE 100(1), 283–299 (2012)

[16] Piètre-Cambacédès, L., Bouissou, M.: Cross-fertilization between safety and security engineering. Reliability Engineering & System Safety, 110–126 (February 2013)

[17] Novak, T., Treytl, A.: Functional safety and system security in automation systems - a life cycle model. In: Proceedings of the IEEE International Conference on Emerging Technologies and Factory Automation (ETFA 2008), pp. 311–318 (September 2008)

Singapore Road Vehicle Fleet Evolution

Wei Wei and Lynette Cheah

Abstract. Vehicle fleet modeling is a useful tool to analyze the dynamics of motor vehicles and their environmental impact at a macroscopic level, and has been applied in the USA and Europe. In this article, a road fleet model is constructed for the city-state Singapore. Policies that control vehicle ownership and congestion road pricing employed since 1998 differentiate Singapore's vehicle market from other markets, making it a particularly interesting case to investigate. The fleet model is constructed using spreadsheets that track vehicle age, vehicle population, vehicle kilometers travelled, fleet fuel use and greenhouse gas (GHG) emissions. The authors hope that the model can be used as a tool to help stakeholders assess the social and environmental impact of relevant policies like capping vehicle growth, scrappage policy, reducing vehicle mileage and adopting green vehicles.

1 Introduction

In Singapore, transport sector is projected to account for 14.5% of greenhouse gas emissions in year 2020 under a business-as-usual scenario (National Climate Change Secretariat 2012). In terms of energy, road vehicles are responsible for most of domestic transport energy demanded (Asia Pacific Energy Research Centre 2013). Vehicle fleet analysis has been used to investigate road vehicle fleet fuel use and GHG emissions under various policy scenarios in the US (Bandivadekar et al. 2008), Europe (Bodek and Heywood 2008; Brand 2010) as well as on a global scale (Facanha et al. 2012), yet it has not been applied to fleets in city-states like Singapore. To understand the current and future environmental impact of Singapore road transportation system, a dynamic road fleet model is constructed to study the timescale of impact of policy changes projected till year 2030.

Wei Wei · Lynette Cheah
Singapore University of Technology and Design,
20 Dover Drive, Singapore, 138682
e-mail: `wei_wei@mymail.sutd.edu.sg, lynette@sutd.edu.sg`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_5

2 Model Structure and Data Source

A Vehicle Quota Scheme (VQS) was introduced in Singapore in May 1990 as a policy instrument to control vehicle ownership (Chin and Smith 1997). Road vehicles must be registered and re-registered under one of the five following categories: cars with engine capacity less than or equal to 1600 cc (category A), cars with engine capacity over 1600 cc (category B), buses, goods vehicles and motorcycles. Potential car buyers must bid for a Certificate of Entitlement (COE) in order to drive on the road. This is tax on top of the cost of purchasing the vehicle. The open bidding process happens twice a month, with a quota pre-allocated by government to restrict vehicle population growth. Therefore, new registration of motor vehicles is equivalent to sales. At the end of year 2013, the road vehicle population numbered 974,170 motor vehicles, among which 66.9% are cars (including taxis), 1.8% are buses, 16.5% are goods vehicles and 14.9% are motorcycles.

An overview of the fleet model's structure is provided in Figure 1. The model is composed of spreadsheets in Microsoft Excel that details vehicle registration and de-registration under VQS, population of different types of motor vehicles, vehicle age distribution and survival rate, vehicle kilometers traveled (VKT), fuel consumption and GHG emissions. Each module of the fleet model is discussed in detail in the next section.

The Singapore Land Transport Authority (LTA) publishes data annually on the city-nation's transportation system, and is the primary source of data for the model. Historical data from 1998 to 2013 is used for the model. Due to limitations in the age distribution data, the model does not include taxis and tax-exempted vehicles. The model stabilizes after year 2008.

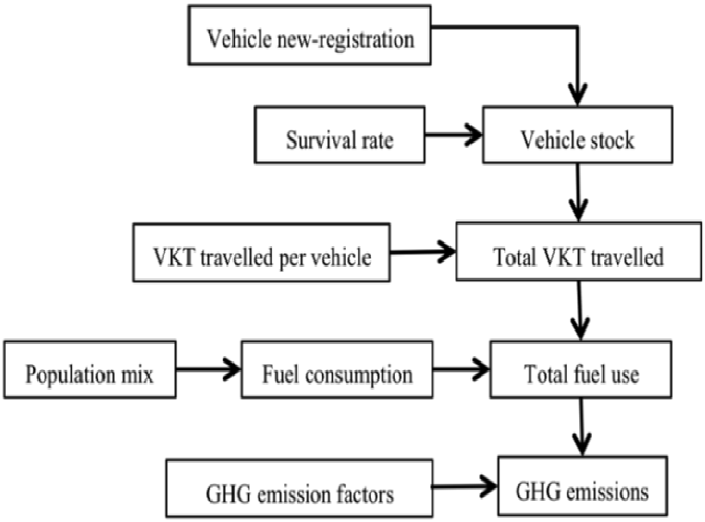


Fig. 1 Overview of Fleet Model Structure

3 Building Blocks of Fleet Model

3.1 *New Registration of Motor Vehicles*

The annual new registration and de-registration of five categories of road vehicles under VQS is shown in Figure 2. New-registration growth rate of buses and motorcycles have not changed much over the past 16 years, with that of goods vehicles has declined from 2005 to 2010, followed by a slight rebound. New car registration peaked at 116,741 cars in year 2006, and has been declining since due to strict government control. As shown in Figure 2, vehicle quota is imposed based on vehicle de-registration in the preceding months.

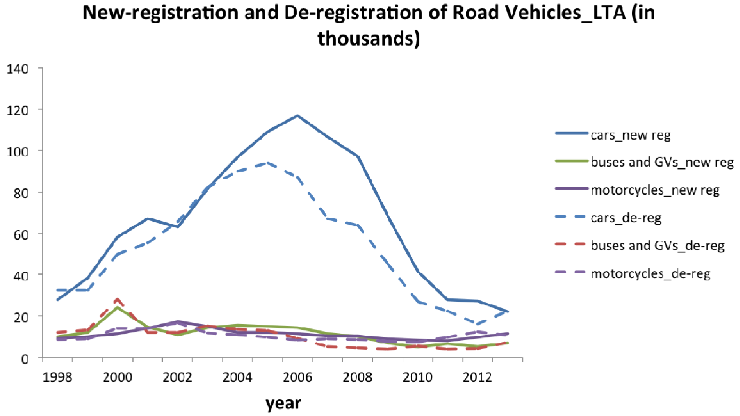


Fig. 2 New Registration and De-registration of Motor Vehicles 1998 -2013 (in thousands)

3.2 *Survival Rate and Lifetime Distribution*

Age distribution of motor vehicles is made available by LTA from year 1998 to 2013 for four vehicle categories: cars (excluding taxis), buses, goods vehicles and motorcycles. Based on this data, we can infer the lifetime distribution of vehicles registered in each year and estimate the survival rate of motor vehicles. Survival rate of vehicles in a given model year is defined as the fraction of that vehicles remaining on the road. Detailed methodology of finding a lifetime distribution that fits LTA data is discussed below.

3.2.1 Cars

In Singapore, the car survival rate is highly influenced by the scrappage policy. For the years studied in the model, COE is only effective for 10 years. Vehicle owners can choose to either renew their COE for another 5 or 10 years by paying the prevailing quota premium for the corresponding year, or de-register their vehicles and have them scrapped or exported. Considering the soaring COE price in recent years, most owners choose to de-register the vehicle at or before 10 years after registration.

The car survival rate estimated from LTA's data, and preliminarily used in the fleet model is shown in Figure 3.

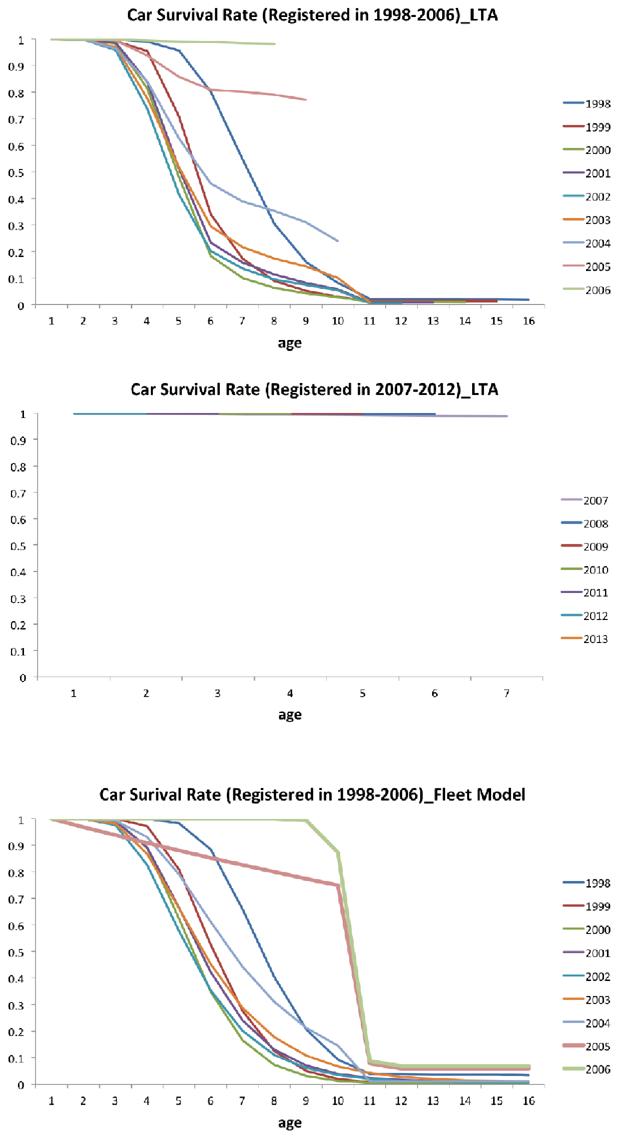


Fig. 3 Car Survival Rate 1998 – 2013 (LTA and Fleet Model)

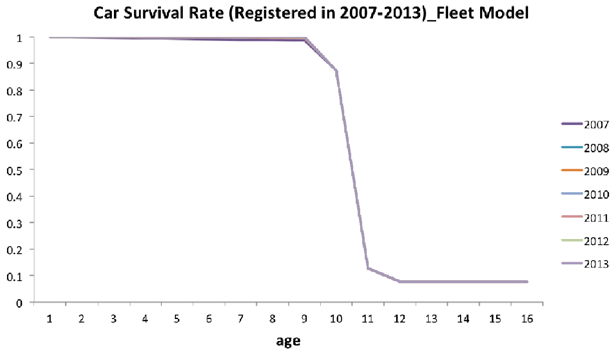


Fig. 3 (*continued*)

In the literature, a logistic model is argued to better describe the removal of vehicles from US fleet, as opposed to an exponential curve as it captures the rapid removal of old cars in the fleet (Field et al. 2000). Bandivadekar et al. likewise use a logistics curve to estimate the scrappage rate of light-duty vehicles in US (2008). Melo discusses three models: normal model, Weibull model and beta model in modeling the disposal of products containing aluminum (1999). It is suggested that the Weibull distribution is the most commonly used model for lifetime distribution, while the beta distribution is more flexible and can model a wide range of distribution shapes. Kolli et al. have fitted beta, gamma, lognormal and Weibull distribution to light car age distribution in France and found Beta and Weibull to be the best fit (2010).

Car survival rate has been quite erratic for the past 16 years. Therefore it is necessary to examine survival rate for each registration year separately in this model. For the Singapore car fleet, logistics curve does not approximate survival rate well as we adjust the parameters. On the other hand, lifetime distribution of cars is found to be suitable for estimating a distribution. Among the four distributions tested: beta, gamma, lognormal and Weibull, lognormal appears to be the best model for year 1998-2004 and year 2006. For cars registered in year 2005 and 2007 onwards, a decay model with constant decay rate is more suitable to estimate car survival rate. The sharp removal of cars at age 11 after year 2005 cannot be modeled by any distribution above. In this case, we assume that the percentage of vehicles removed at age 11 from year 2004 onwards is the same as that of 2003. The removal of cars after age 11 is also assumed to be at a constant rate of 1-100 vehicles per year. In Singapore, cars have a median lifetime of around 5 years only.

Table 1 Car Lifetime Distribution Parameters

1998	1999	2000	2001	2002	2003	2004	2005
Lognormal	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal	Lognormal	Decay
$\mu = 1.8883$ $\sigma = 0.2339$	$\mu = 1.6271$ $\sigma = 0.2758$	$\mu = 1.4857$ $\sigma = 0.3163$	$\mu = 1.5345$ $\sigma = 0.3554$	$\mu = 1.4602$ $\sigma = 0.3877$	$\mu = 1.5556$ $\sigma = 0.4110$	$\mu = 1.7204$ $\sigma = 0.4214$	$\beta = 0.9685$

2006	2007	2008	2009	2010	2011	2012	2013
Lognormal	Decay	Decay	Decay	Decay	Decay	Decay	Decay
$\mu = 2.2874$ $\sigma = 0.0825$	$\beta = 0.9983$	$\beta = 0.9996$	$\beta = 0.9996$	$\beta = 0.9996$	$\beta = 0.9996$	$\beta = 0.9996$	$\beta = 0.9996$

For Singapore, the deviation from previous literature of modeling lifetime distribution as beta, Weibull or survival rate as logistics curve is speculated to be a result of government control of vehicle ownership. A PARF (Preferential Additional Registration Fee) rebate and a COE rebate are the two rebate schemes to encourage early car de-registration. Essentially, COE affects both vehicle sales directly and de-registration indirectly. The decision made by car owners can be considered as a trade-off between the rebate and Quota Premium required to pay to buy a new car. In previous studies, a life cycle cost analysis for vehicles in Singapore has been performed (Wong et al. 2010). In Figure 4 below, we visualize the timescale of this relationship for a Toyota Corolla registered in year 2002 with an Open Market Value (OMV) of $19,384 and a Quota Premium of $ 32,981. The COE rebate decreases linearly with time, while PARF rebate is tiered. Quota premium reached its lowest during 2008's economic crisis and has been soaring in recent years. It is reasonable for car owners to keep their cars while the aggregated value of rebate and quota premium is positive. It should be noted that at each given point in time, car owners do not know the quota premium in the future.

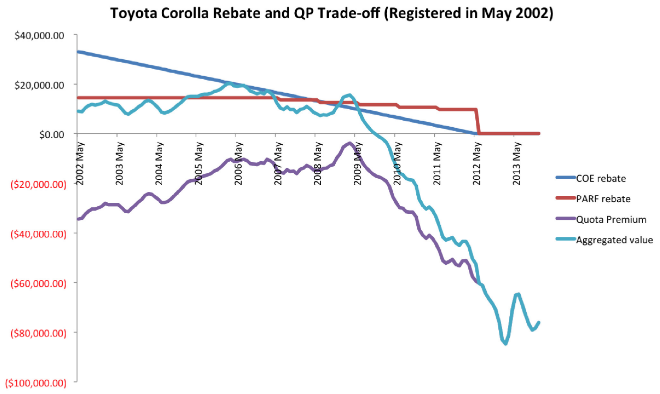


Fig. 4 Illustrative Example of Trade-off between Rebate and QP for Cars

3.2.2 Buses

For buses, there are some uncertainties in different registration years' survival rate. A common pattern is that around 80% of buses remain in the fleet before COE expires at 10 years after registration, with a steeper slope at the 10-year point representing the influence of the scrappage policy. For simplicity, we categorize registration years into two groups based on population decay rate and use two decay models to describe vehicle population change. For time periods 1998 - 2001 and 2004 - 2030, we use the same decay rate as year 1998; for time period 2002 – 2003 we use the same decay rate as year 2002 in LTA data (Figure 5). The result of this assumption appears to be coherent with LTA vehicle population data shown in Figure 5.

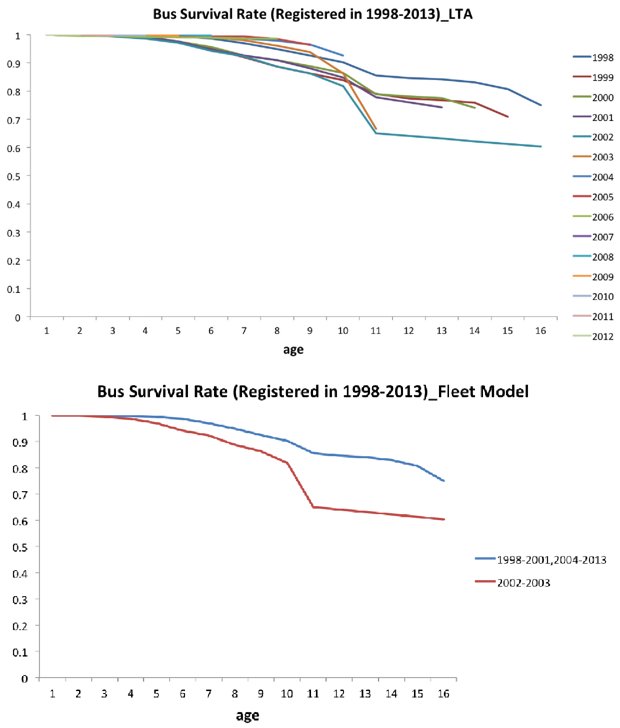


Fig. 5 Bus Survival Rate 1998 – 2013 (LTA and Fleet Model)

3.2.3 Goods Vehicles

There is very high variability in goods vehicles' survival rate, especially for years 1998-2003. To ensure the accuracy of the model, similar with bus population, we separate the registration years into 3 groups based on their similarities in population decay rate: 1998-1999, 2000-2002 and 2003 onwards. The result is shown in Figure 6.

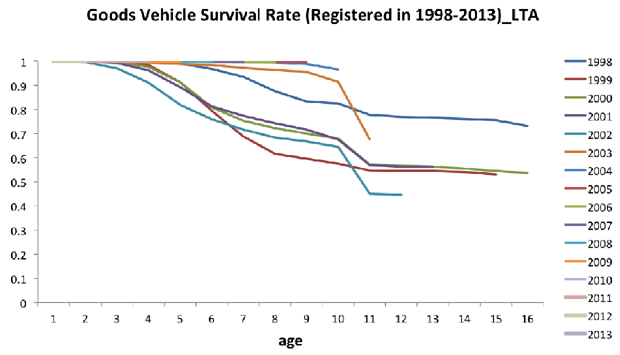


Fig. 6 Goods Vehicles Survival Rate 1998 – 2013 (LTA and Fleet Model)

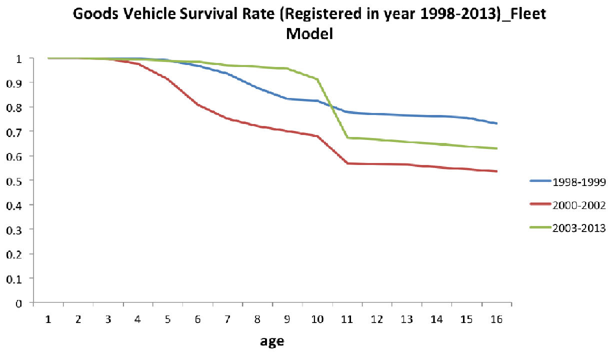


Fig. 6 (*continued*)

3.2.4 Motorcycles

Motorcycles' survival rate behaves reasonably consistent so we use the same population decay rate as year 2002 in LTA data for all registration years in the fleet model. The result is shown in Figure 7.

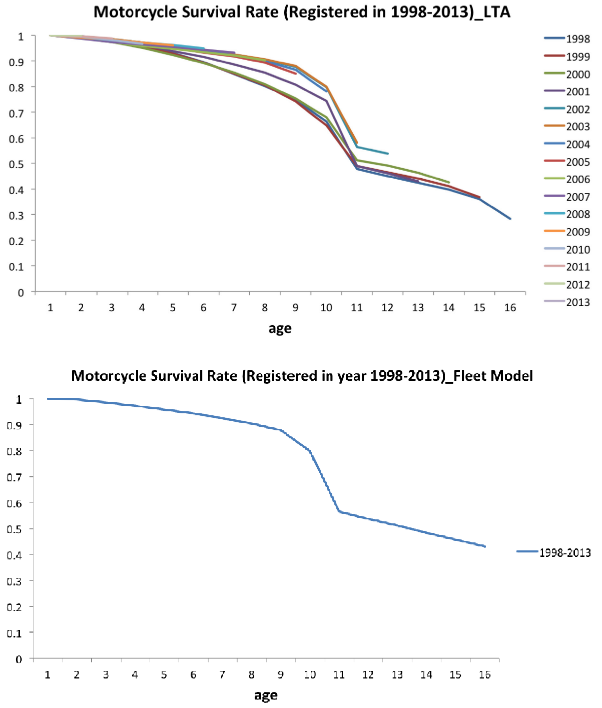


Fig. 7 Motorcycle Survival Rate 1998 – 2013 (LTA and Fleet Model)

3.3 Vehicle Kilometers Travelled (VKT)

The model uses the following equation to calculate total VKT in year j:

$$VKT_j = \sum_i N_{i,j} \times VKT_{i,j} \quad (1)$$

Where $N_{i,j}$ denotes the number of vehicles registered in year i remained in calendar year j's stock, $VKT_{i,j}$ denotes the average per-vehicle kilometers travelled in calendar year j for vehicles registered in year i.

As average per-vehicle kilometers travelled data from LTA is only available from year 2003 to 2012, the model assumes the same values for year 1998-2002 as 2003 data and for year 2013 as 2012 data. Since the data does not distinguish vehicles registered in different years, we assume each registration year's vehicle model behaves in the same way, though this may not be the case and more data collection is needed to account for the difference. Figure 8 shows total VKT for four vehicle categories calculated using the equation above.

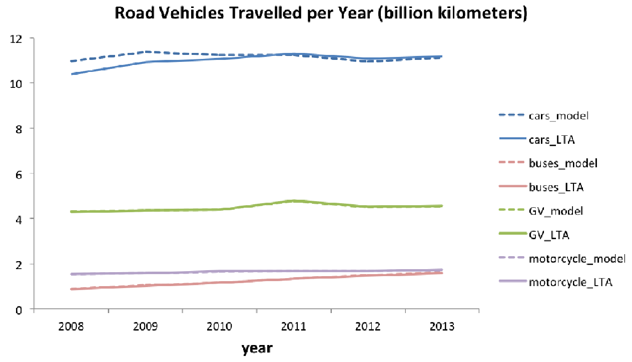


Fig. 8 Motor Vehicles Travelled per Year in Billion Kilometers (LTA and Fleet Model)

3.4 Vehicle Fuel Use

A Fuel Economy Labeling Scheme (FELS) was first launched as a voluntary program in Singapore in 2003, it then become mandatory for all car retailers to display fuel economy labels in their showrooms since 2009. The fuel economy data provided by LTA only includes year 2003, 2004, 2007 and 2008 for Category A cars (≤1600cc), Category B cars (>1600cc), light goods vehicles (LGV), heavy goods vehicles (HGV) and very heavy goods vehicles (VHGV) (excluding year 2003) (See Table 2).

Table 2 Fuel Consumption Data (Land Transport Authority 2005; Land Transport Authority 2008; Mercedes-Benz 2014)

Fuel Consumption (liters/100km)					
From LTA Statistics Brief					From Mercedes-Benz Website
	2003	2004	2007	2008	2012
Category A Car	8.3	10.0	9.6	9.4	-
Category B Car	12.5	12.9	12.0	11.5	-
LGV	10.0	10.0	10.1	10.0	-
HGV	12.5	11.9	14.7	14.9	-
VHGV	-	47	50.5	50.2	-
Buses	-				42.31

To speculate fuel consumption data for the missing years, we investigate the top 8 car makes sold in Singapore, which constitutes 80% of total car population on average. From Figure 9, the car mix has not experienced a big change over the past 11 years. Fuel economy data from US-based fueleconomy.gov also suggests that different model yeas have quite consistent fuel economy performance for popular car models like Toyota Corolla, Toyota Camry for model year 2003-2008, Nissan Sentra for model year 2003-2006, and Mercedes SLK, CLK series for model year 2003-2011 (US Department of Energy et al. 2014). The fleet model takes these factors into account and presents an estimation of the fuel consumption performance for cars.

For goods vehicles, Toyota, Nissan, Mitsubishi and Isuzu account for on average 73.5% of total population for year 2003 – 2013. Mercedes Benz and Toyota are the top two car makes for bus category. Due to the lack of fuel consumption data for these makes, we estimate the figure for goods vehicles from LTA's data in Table 2 and for buses from Mercedes Benz bus models. As for motorcycles, Honda and Yamaha are the two dominant makes in Singapore, occupying 72.7% of total population for the past 11 years. The fuel consumption for motorcycle is estimated by taking the average of Honda and Yamaha models from totalmotorcycle.com (2014).

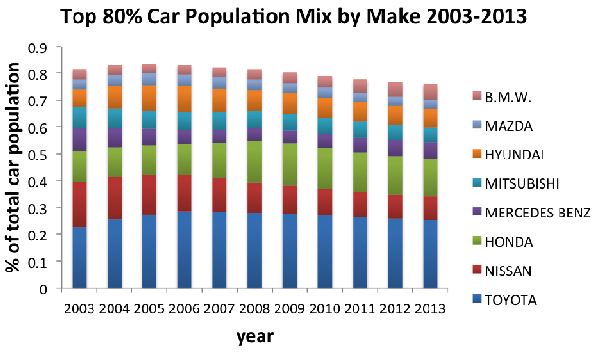


Fig. 9 Car Population by Make (Land Transport Authority 2013b)

Figure 10 shows the resultant vehicle fuel use from year 2008 to 2013 calculated using the following equation:

$$F_j = \sum_i N_{i,j} \times VKT_{i,j} \times f_{i,j} \tag{2}$$

Where $N_{i,j}$ and $VKT_{i,j}$ have the same meaning as that equation (1) and $f_{i,j}$ denotes the fuel consumption in calendar year j for vehicles registered in year i.

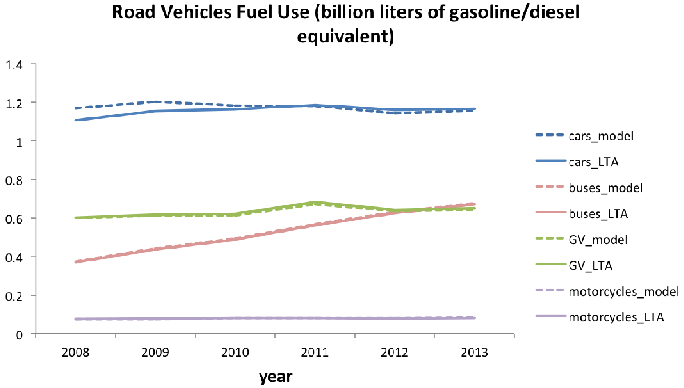


Fig. 10 Motor Vehicles Fuel Consumption in Billion Liters (LTA and Fleet Model)

The model assumes that cars with engine capacity less than or equal to 1600 cc (Category A) has the same average per-vehicle VKT as cars with engine capacity larger than 1600 cc (Category B), so total VKT for cars is split into the two car categories using the population ratio.

3.5 Greenhouse Gas (GHG) Emissions

Greenhouse gas emission is calculated by multiplying fuel use result obtained in 3.4 by GHG emission factors provided by Carbon Trust (2013). 99% of cars and motorcycles use gasoline from 2003 to 2013, while around 93% of buses and goods vehicles use diesel. The model thus treats all cars and motorcycles as gasoline-powered, and all buses and goods vehicles as diesel-powered. Detailed data is shown in Table 3.

Table 3 GHG Emissions Factors (Carbon Trust 2013)

GHG Emission Factors (kgCO_2/liter)				
	Cars	Buses	Goods Vehicles	Motorcycles
Gasoline	2.2144	-	-	2.2144
Diesel	-	2.6008	2.6008	-

4 Model Findings

The projection of the fleet model till year 2030 is then evaluated under three scenarios: a baseline scenario with 0.5% vehicle new registration growth rate, scenario I with 5% vehicle new registration growth rate and scenario II with 10% vehicle new registration growth rate. The projected vehicle stock of cars, buses, goods vehicles and motorcycles is included in Figure 11.

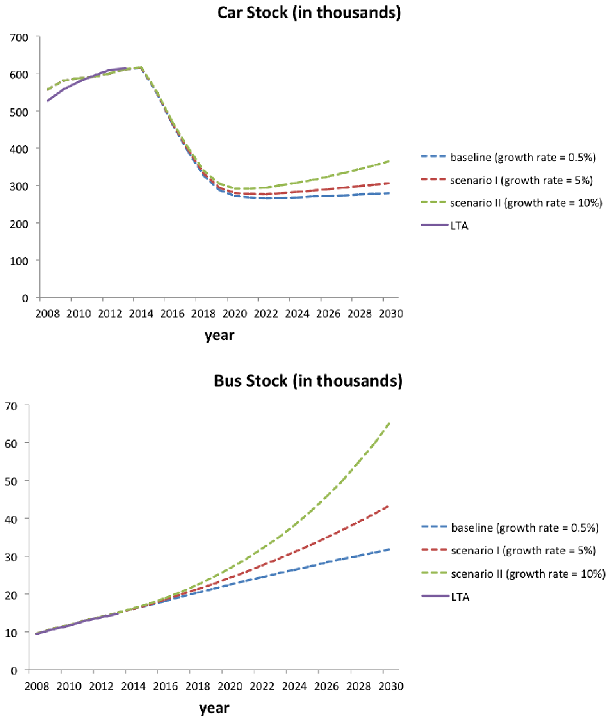


Fig. 11 Model Projected Vehicle Stock in Thousands till Year 2030

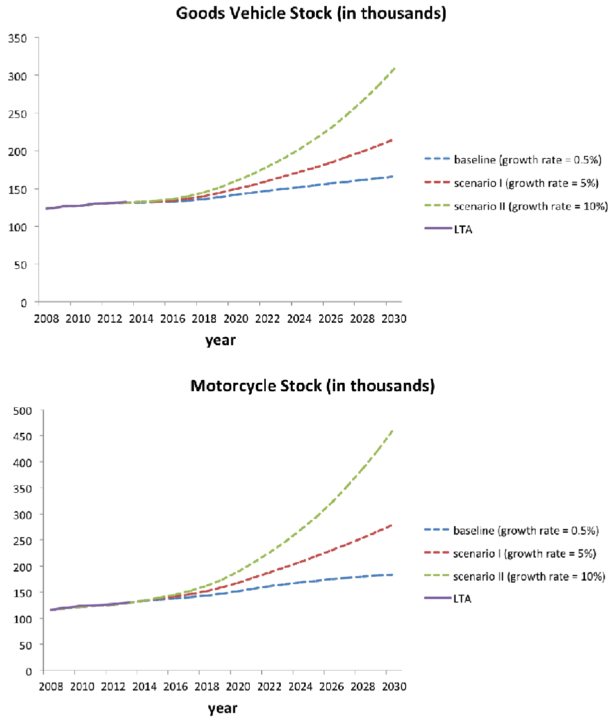


Fig. 11 (*continued*)

It is observed that car population will decrease from year 2014 to 2020 even at an annual 10% vehicle new registration growth rate. This is attributed to the high new registration of cars in 2004-2008. The removal of a large percentage of vehicles when a bulk of existing COE expires after 10 years causes this drop in car stock. The new registration and de-registration of motor vehicles from the model are shown in Figure 12. For buses, goods vehicles and motorcycles, there is no such phenomena and the increase in new-registration growth rate leads to increase in vehicle stock. VKT and fuel consumption roughly follow the same distribution as vehicle stock, as there is no big change in VKT per vehicle or fuel economy in the model. The result for cars is shown in Figure 13.

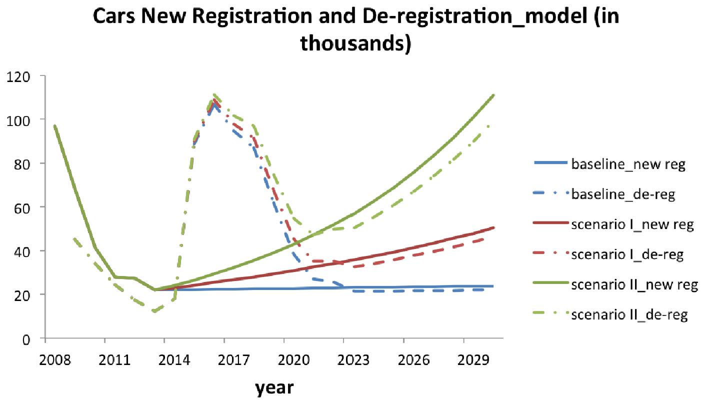


Fig. 12 Cars New registration and De-registration from Fleet Model (in thousands)

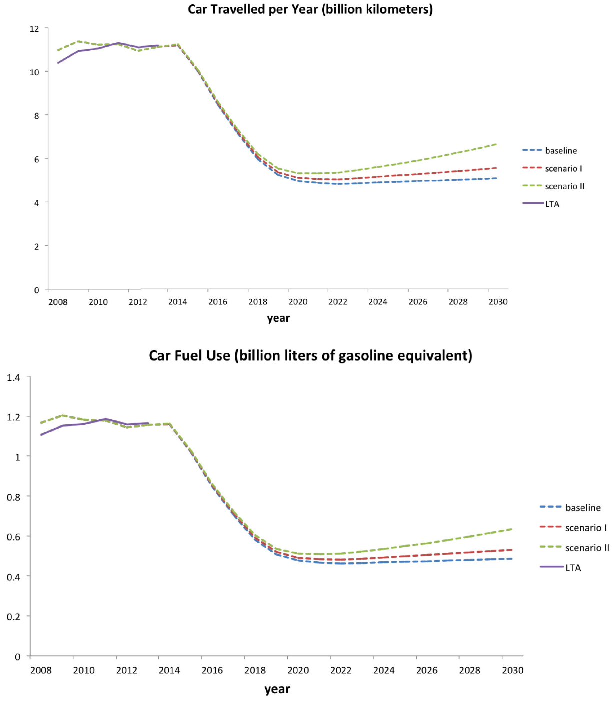


Fig. 13 Car Projected VKT and Fuel Use till year 2030

The relative vehicle stock, VKT, fuel consumption and GHG emissions of cars, buses, goods vehicles and motorcycles under the baseline scenario is shown in Figure 14. It can be observed that the decrease in car stock after year 2014 makes the VKT of cars and goods vehicles roughly the same after year 2020. Total fuel consumption of road fleet is mainly attributed to cars for time period 1998-2014, then to buses for year 2014 onwards.

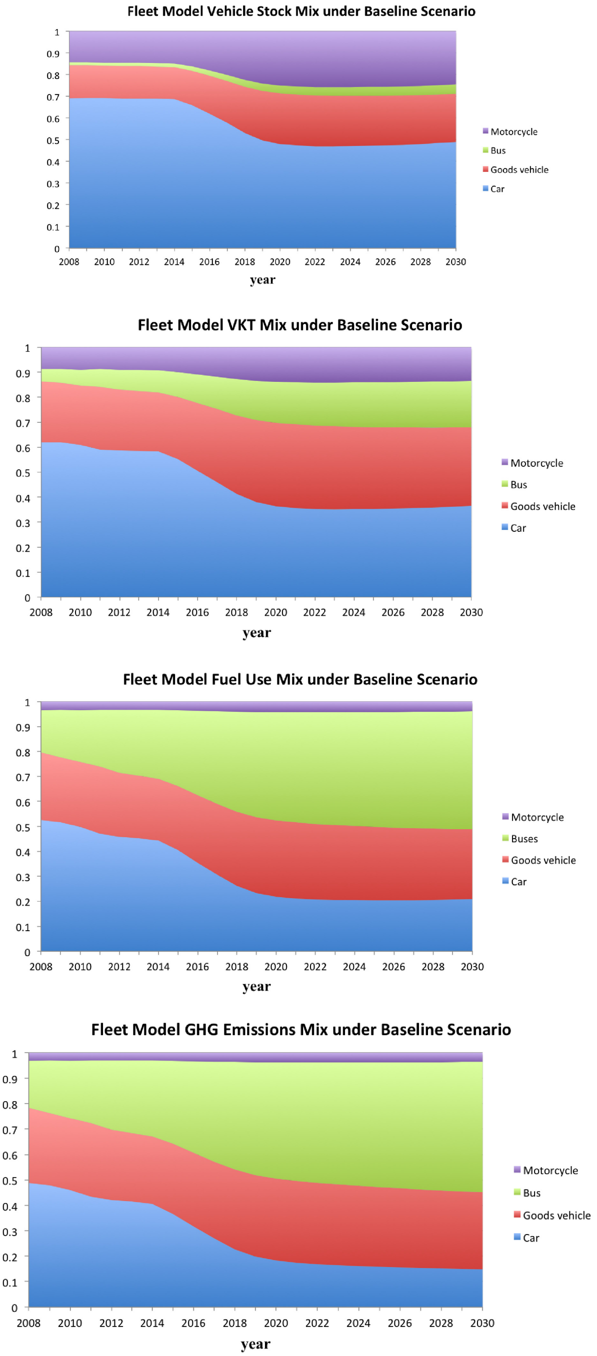


Fig. 14 Fleet Model Vehicle Stock, VKT, Fuel Use and GHG Emissions Mix under Baseline Scenario

5 Future Work

Using the fleet model, a sensitivity analysis will be performed to evaluate the impact of different input parameters including scrappage policy, adoption of green vehicles, VKT change on vehicle population and fuel consumption. The energy outlook and emission target for year 2020 can also be compared with the model result under different scenarios.

6 Conclusions

The paper presents the construction of a road fleet model for Singapore and how it can be used to analyze vehicle stock, vehicle kilometers travelled, fuel consumption and GHG emissions of Singapore road vehicles. A lognormal and decay model is found to best describe the lifetime distribution for cars, while for buses, goods vehicles and motorcycles, decay model is a good approximation. Singapore road fleet is highly influenced by government policies, thus differentiating its behavior from other countries. Based on the model's projection till year 2030, car stock is predicted to decrease from year 2014 to 2020 even when vehicle new-registration annual growth rate is kept at 10% from 2014. This also leads to the dominating role of buses in road transport fuel use after year 2014. The authors hope that this road fleet model can be used as a tool to assess policy impact on road transport system in Singapore.

Acknowledgements. The authors would like to thank Land Transport Authority (LTA) for its kind assistance in providing data for this study.

References

Asia Pacific Energy Research Centre, APEC Energy Demand and Supply Outlook, 5th edn., p. 424 (2013)

Bandivadekar, A., Bodek, K., Cheah, L., et al.: On the Road in 2035: Reducing Transportation's Petroleum Consumption and GHG Emissions, Cambridge (2008)

Bodek, K., Heywood, J.: Europe's Evolving Passenger Vehicle Fleet: Fuel Use and GHG Emissions Scenarios through 2035. MIT Lab. Energy Environ. (2008)

Brand, C.: UK Transport Carbon Model: Reference Guide, Oxford (2010)

Carbon Trust, Conversion Factors (2013), `http://www.carbontrust.com/media/18223/ctl153_conversion_factors.pdf` (accessed May 4, 2014)

Chin, A., Smith, P.: Automobile ownership and government policy: The economics of Singapore's vehicle quota scheme. Transp. Res. Part A Policy Pract. 31, 129–140 (1997), doi:10.1016/S0965-8564(96)00012-2

Facanha, C., Blumberg, K., Miller, J.: Global Transportation Energy and Climate Roadmap, Washington (2012)

Field, F., Kirchain, R., Clark, J.: Life-Cycle Assessment and Temporal Distributions of Emissions Developing a Fleet-Based Analysis. J. Ind. Ecol. 4, 71–91 (2000), doi:10.1162/108819800569816

Kolli, Z., Hivert, L., Dupont-Kieffer, A.: Car survival in a national fleet: a non-parametric approach based on French data. In: 12th World Conf. Transp. Res., Lisbonne, pp. 1–17 (2010)

Land Transport Authority, Land Transport Master Plan 2013, Singapore (2013a)

Land Transport Authority, Singapore Land Transport: Statistics in Brief 2005, Singapore (2005)

Land Transport Authority, Singapore Land Transport: Statistics in Brief 2008, Singapore (2008)

Land Transport Authority, Annual Vehicle Statistics: Total Car Population by Make, Singapore (2013b)

Melo, M.T.: Statistical analysis of metal scrap generation: the case of aluminium in Germany. Resour. Conserv. Recycl. 26, 91–113 (1999),
doi:10.1016/S0921-3449(98)00077-9

Mercedes-Benz, Mercedes-Benz Buses Fuel Consumption (2014),
`http://www2.mercedes-benz.co.uk/content/unitedkingdom/mpc/mpc_unitedkingdom_website/en/home_mpc/bus/home/buses_world/record_run/result/Promise_redeemed.html` (accessed May 4, 2014)

National Climate Change Secretariat, Climate Change & Singapore: Challenges. Opportunities. Partnerships (2012)

Total Motorcycle. Fuel Economy Guide (2014),
`http://www.totalmotorcycle.com/` (accessed April 16, 2014)

US Department of Energy, Energy Efficiency & Renewable Energy, Office of Transportation and Air Quality, US Environmental Protection Agency. Fuel Economy (2014), `http://www.fueleconomy.gov/` (accessed April 26, 2014)

Wong, Y.S., Lu, W., Wang, Z.: Life Cycle Cost Analysis of Different Vehicle Technologies in Singapore. World Electr. Veh. J. 4, 912–920 (2010)

Design and Management of Flexible On-Shore LNG Production Engineering Systems

Michel-Alexandre Cardin, Mehdi Ranjbar-Bourani, and Richard de Neufville

Abstract. This paper presents a flexibility analysis as a practical procedure to evaluate large-scale capital-intensive projects considering market uncertainty. It considers the combined effects of the time value of money, economies of scale, and learning, and demonstrates the additional benefits stemming from considerations of uncertainty and flexibility in the early stages of design and project evaluation. This study focuses on the long-term deployment of liquefied natural gas (LNG) technology in a target market to supply the transportation sectors. Two design alternatives are considered: 1) fixed design, a big centralized production facility; 2) flexible modular designs, either using phasing approach at the big plant site or the same flexible approach with an option to move modular plants at distance. To compare the design alternatives, a structured flexibility methodology is applied based on several economic lifecycle performance indicators (e.g. Net Present Value, Initial CAPEX, etc.). Results indicate that a flexible modular deployment strategy improves the economic performance as compared to optimum fixed designs. They also indicate that factoring flexibility to locate modules at a distance further improves system performance. Such improvement enhances as learning rate increases. Overall, the study shows that flexibility in engineering design has multiple, supporting advantages due to uncertainty, location and learning.

1 Introduction

The advantage of using natural gas products has increased over the last three decades, resulting in a considerable demand growth for LNG. Research has shown

---

Michel-Alexandre Cardin · Mehdi Ranjbar-Bourani
Department of Industrial and Systems Engineering, National University of Singapore - Block E1A #06-25, 1 Engineering Drive 2, 117576, Singapore

Richard de Neufville
Engineering Systems Division, Massachusetts Institute of Technology – 77 Massachusetts Avenue, Cambridge, MA, 02139, United States

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_6

that by 2030 there is a possibility that the overall LNG demand will be more than three times higher than from where it was in 2011 and the regional distribution will significantly change accordingly [1]. More specifically, gas product demand and supply forecasts in a target market indicate a potential shortfall of 300 to 600 TJ/day by 2015, and between zero and 600 TJ/day by 2020 [2]. A combination of growth and replacement production indicates there is a need to source at least 1,100 TJ/day of new production by 2020.

Since LNG can be used reliably as on-road transport fuel, there are growing business opportunities for LNG production. Development of this business can be risky, however, as it requires substantial amount of initial investment. The project will be subject to uncertainty in LNG demand, gas price, and facility availability. The design stage of such projects is significantly large as critical decisions need to be made, and as changing the system configuration later on might be too costly.

This study presents flexibility analysis as a practical procedure to maximize the expected value of a system over its useful time. It enables developers to adapt the system for better performance as its requirements and opportunities evolve over its useful life by exploiting the notion of modularity in design [3,4]. The study contrasts and compares to others as it considers explicitly the combined effects of uncertainty, the time value of money, economies of scale (EoS), and learning to highlight the economic benefits stemming from flexibility. It is first to do this in the context of LNG production systems.

The rest of this paper is organized as follows. Section 2 discusses the motivations to apply the practical flexibility procedure, which considers explicitly uncertainty and flexibility, to designing and evaluating LNG production systems. It also reviews relevant literature. Section 3 identifies the research gap, and defines the scope of the problem under consideration. Section 4 describes the methodology in generic terms. Section 5 presents a case study on a LNG production system that demonstrates the implementation of the analytical approach. Section 0 summarizes major findings, providing conclusions and insights for further research.

2 Background and Motivation

2.1 Flexibility in Engineering Design

Flexibility in engineering design is an interdisciplinary field for research and practice [3,4]. It adapts the concept of financial options to real engineering systems, with the goal of increasing the expected economic value by providing the “right, but not the obligation to change a system” to respond to uncertainties most profitably [5]. Flexibility exists “on” and “in” engineering systems. Flexibility “on” systems is associated with managerial flexibility like abandoning, deferring until favorable market conditions, expanding/contracting/reducing capacity, deploying capacity over time, switching inputs/outputs, and/or mixing the above [5]. Flexibility “in” systems refers to technical engineering and design components enabling real options – another word for flexibility – in deployment and operations [6].

Cardin [4] provides a taxonomy and design framework to organize design and evaluation activities to enable flexibility in engineering systems.

Flexibility enables a system to capture the potential value associated with different scenarios. It enables, for instance, capturing more demand in high demand cases, thus increasing the expected economic value (i.e. like a call option). It might reduce financial losses in a downside demand scenario (i.e. like insurance).

2.2 Simulation Based Flexibility Analysis

Monte Carlo simulation combined with an approach base on managerial decision rules is used to simulate the behavior of systems in different applications. A decision rule is a triggering mechanism based on a criterion – typically related to an observed state of uncertainty – determining the appropriate moment to exercise the flexibility in operations. This method is now widely accepted for evaluation of flexibility in engineering design [3,4]. The rationale for using this method emerges from the fact that using theoretical methods from finance have serious shortcomings, especially for solving complex real-world problems. On the other hand, Monte Carlo simulation provides a platform so that even a complex system can be modeled easily. Theoretical evaluation methods relying on standard real options analysis (e.g. binomial lattice) used for complex projects over simplify the original problem so that it can be solved. These simplifications can lead to inaccurate results. By using Monte Carlo simulation and decision rules, one has the freedom to incorporate precisely the detailed attributes of the real-world problem by parameterizing the physical design variables, analytical parameters and decision rules. Fitzgerald et al. [7] presented an extended version of a simulation based analysis called Epoch Era Analysis (EEA) [8,9] to investigate the value of changeability in complex engineering systems at early stage of the design process. They used transition rule matrix in EEA whereas in this study different decision rules and their corresponding parameters were used in the Monte Carlo Simulation framework.

2.3 LNG Production System Design

LNG production system design has become more critical due to the growth of natural gas supply and demand and the great risks in this industry. Literature has shown a growing research towards designing value LNG production systems focusing on different segments of the LNG supply chain, depending on the problem under consideration and geographical situation. Özelkan et al. [10] studied the coupled segments of large scale shipping and receiving terminal of an LNG supply chain to minimize cost and storage inventory, while maximizing the output of natural gas to be sold to the market. Grønhaug, Christiansen [11] presented both an arc-flow and a path-flow model for tactical planning to optimize the LNG inventory routing problem. Andersson et al. [12] worked on transportation planning and inventory management of a LNG supply chain used in tactical planning during negotiations about deliveries to different regasification terminals and annual delivery plan used in operational level decision making.

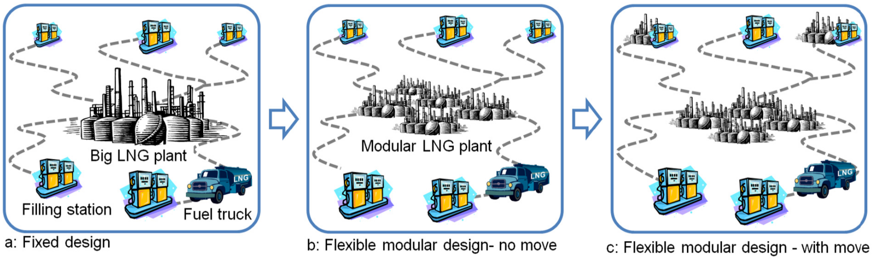
As the overview suggests, more work is needed to evaluate LNG production systems in the early stages of design. In particular, more efforts are needed considering strategic level decisions involving flexibility and uncertainty in the analysis of site production capacity and deployment over time. In addition, to these authors' knowledge there has been no other study considering the combined effects of economies of scale, time value of money, and learning in this context.

The main contribution of the paper is to investigate these effects on key strategic factors affecting the design of LNG production systems, from onshore natural gas transmission pipeline to end users at candidate geographical demand sites. The goal is to identify designs and decision rules that provide better expected economic value over the entire lifetime of a project, as compared from the typical outputs from standard design and project evaluation.

3 Scope and Problem Definition

This study focuses on the design and development of the LNG production system to provide fuel for trucks used in on-road product transportation in a target market. The goal is to meet the LNG demand at different geographical sites, knowing that these sites have direct access to an existing natural gas pipeline. Figure 1 schematically represents the LNG production system, from a fixed towards a more flexible design. This example has five candidate demand points equipped with filling station facilities and a main production site dedicated to a centralized LNG plant. All sites have access to the on-shore pipeline distributing the natural gas. In the main production site, LNG produced through the liquefaction process is transferred to the candidate demand sites.

Two main LNG system designs are investigated: 1) fixed centralized design (a), and 2) flexible modular designs, (b and c). In the fixed centralized design, the optimal capacity significantly depends on the strength of the economies of scale. A big LNG plant is built in the main production site and LNG produced is carried to the market sites using fuel trucks. The flexible modular designs includes: 1) flexible modular design – no move, (b), which considers a phasing approach using a modular LNG plant with the flexibility to expand capacity at the main production site, and transport LNG to demand sites; 2) flexible modular design with move, (c), which is the same design as the flexible modular design, (b),but with the ability to move the modular LNG plants to other demand sites.



a: Fixed design
b: Flexible modular design- no move
c: Flexible modular design - with move

Fig. 1 Shift from a fixed LNG system (a) design towards more flexible systems (b, c)

4 Methodology

This paper uses a practical four-step process to quantify flexibility under uncertainty [3]. This approach improves the lifecycle performance of a project dependent on a range of potential uncertainties. Several economic lifecycle performance indicators are used (e.g. Net Present Value, Initial CAPEX, etc.) to quantify the "Value of Flexibility". The steps below describe the generic process followed as it relates to LNG demand growth.

4.1 Step 1: Deterministic Analysis

The proposed methodology starts with a deterministic analysis. The aim is to understand the key components of the system that influence its lifecycle performance. The performance metric used in this problem is NPV, calculated as the sum of discounted cash flows throughout the project lifecycle T = 20 years – see Equation (1). Variables TR_t and TC_t are the total revenues and costs incurred in years t = 1, 2, ... T, and r is the discount rate.

$$NPV = \sum_{t=1}^{T} \frac{TR_t - TC_t}{(1+r)^t} \tag{1}$$

LNG demand is a key driver of system performance. A deterministic S-curve functionis assumed to simulate LNG demand at time t (D_t) over the study period, as shown in Equation (2). The rationale is that LNG demand initially grows slowly; it then increases exponentially, and finally tapers as it approaches a saturation limit. Variable M_T is the maximum expected demand for LNG, b is the sharpness parameter that determines how fast demand grows over time to reach the upper bound for demand. The parameter a translates the curve horizontally.

$$D_t = \frac{M_T}{1 + ae^{-bt}} \tag{2}$$

where a is calculated using Equation (3).

$$a = \frac{M_T}{D_0} - 1 \tag{3}$$

In general, the conventional DCF model is built to assess the performance of the system under deterministic conditions. This step captures standard industry practice in terms of design and project evaluation [13].

4.2 Step 2: Uncertainty Analysis

The analysis under uncertainty considers a distribution of outcomes instead of a single performance output. Hence, in this step NPV_S, which refers to NPV under demand scenario *s*, is calculated in terms of different realized demand scenarios. A stochastic S-curve function simulated LNG demand over the system's lifecycle using additional uncertainty factors, as shown in Equation (4).

$$R_t = \frac{M_T \pm \Delta_{M_T}}{1 + a_u e^{-(b \pm \Delta_b)t}} \tag{4}$$

Simulation is used to simulate a wide range of LNG demand scenarios. This analysis recognizing uncertainty provides designers a more realistic overview of system performance as compared to the deterministic analysis in Step 1.

4.3 Step 3: Flexibility Analysis

To account for system flexibility, decision rules are embedded into the DCF model under uncertainty. For example, to embed the capacity expansion policy in flexible modular designs, a simple decision rule is programmed in the Excel® spreadsheet DCF model under uncertainty. For instance a capacity expansion policy can be: IF "*observed aggregate demand in the current year is higher than a certain threshold value at the main production site*" THEN "*build extra modular plant as capacity expansion policy"* ELSE *"do nothing*". The threshold value determines when extra capacity should be built, either at the main production site or other demand sites. For example, decision-makers may decide to add another modular plant as soon as the difference between the realized and current capacity (i.e. unmet demand) reaches 60% of the capacity of a modular plant for the site. The value of flexibility is calculated as shown in Equation (5).

$$\text{Flexibility Value} = \max\left(0, \text{ENPV}_{\text{Flexible design}} - \text{ENPV}_{\text{Optimum fixed design}}\right) \tag{5}$$

4.3.1 Multi-criteria Decision Making Table

To evaluate flexible designs, the analyst needs to factor in a distribution of outcomes instead of one single point to support design decision-making. These distributions can be interpreted using the shape of different criteria. For instance, one may seek to maximize ENPV or to minimize downside risk or to choose some balance between these criteria. Given the several criteria that are not directly compatible, it is useful to create a multi-criteria table, providing decision makers with the information needed to trade-off criteria among flexible design alternatives.

4.4 Step 4: Sensitivity Analysis

A sensitivity analysis is performed to observe how the system responds to different parameters and input data. This study investigated the effects of varying discount rate, economies of scale, and learning on the results. These parameters capture key tradeoffs in engineering design and economic analysis.

5 Application and Discussion

5.1 Modeling Assumptions

The following assumptions are made. Demand is assumed to be evenly distributed in the region over five distinct demand sites. There is no market at the main production site. All sites have access to on-shore natural gas pipeline in the target region. Time to build is 2 years for the big plant, but only 1 year for small plants. Also, if one decides to expand capacity in year t, extra capacity will be available for production in year t+1. The project lifetime is 20 years. Ten-year straight-line depreciation is used for all LNG production facilities with zero salvage value. The discount rate is 10%, and the corporate tax rate is 15%. Parameters associated with deterministic and stochastic LNG demand modeling are summarized in Table 1.

Regarding design parameters, the capacity of modular LNG plant is set to 25 tpd with initial capital expenditure (Capex) $25 million. The Opex of the plant is assumed 5% of the plant's Capex. Flexibility cost is 10% of the Capex of the first capacity deployment at each site because of gas tie-in to the existing natural gas pipeline and extra land cost. Transportation cost is set to $0.4 per ton-kilometer, while travel distances from the main production site to sites 1 to 5 are 118, 121, 281, 318, and 446 Km respectively.

Table 1 Parameters used in uncertainty modeling for each demand site

Parameter	Deterministic demand	Parameter	Stochastic demand
D_0	5 tpd	Δp_0	50%
M_T	50 tpd	Δp_T	50%
a	9	Gt	~ Normal(0,1)
b	0.35	Δ_b	70%
T	20 years	Δ_{av}	5%

5.2 Step 1: Deterministic Analysis

Results show the NPV for different sizes of plants that have various economies of scale factors. It shows, as might be anticipated intuitively, that: a) for any set of plant size and economies of scale, there is a "sweet spot": build too small, and there is no profit from higher demands; build too large, and there is risk of overcapacity and attendant losses, and b) the greater the economies of scale, the larger the fixed design should be. The advantages of these economies compensate for the overcapacity of the greater size over initial demand, and counterbalance the economic advantages of deferring costs (due to the discount rate). Note however, that deterministic analysis based on expected LNG demand may give incorrect results, compared to realistic analysis that recognizes uncertainty, as shown next.

5.3 Step 2: Uncertainty Analysis

The deterministic analysis gives a misleading impression of lower value due to the Flaw of Averages [14]. Engineering systems typically respond non-linearly to inputs, and any decision based on average value of these factors is almost certain to provide a false reading on the actual average value of an alternative. To get the right answer, one needs to analyze the system under uncertainty.

The case study recognizes LNG demand as a key source of uncertainty. Using Monte Carlo simulation, it explores how design alternatives behave under different LNG demand scenarios. Simulations use different LNG plant capacities and economies of scale parameters. The aim is to find the stochastic optimum design for plant capacity. The results show that when using 2,000 demand scenarios, the system performance converges to a steady state value with negligible variations.

Table 2 compares the results of the deterministic and uncertainty analyses. It shows that optimum capacities and values generated by the uncertainty analysis are systematically different (in this case, smaller) than those obtained from the deterministic analysis.

The intuition is that an asymmetric response of the system occurs because of variations in demand: lower demands lead to losses, which higher demands can only partially compensate, because of limitations in installed capacity. This reality favors smaller capacity designs that cost less and minimizes unused capacity when uncertainty is considered, as compared to a deterministic analysis, which typically favors more capacity to be deployed upfront.

Table 2 Optimum fixed designs under deterministic and uncertain LNG demand with different Economies of Scale parameters α

Economies of scale parameter, α	Optimum capacity (ton per day)		Optimum value ($ millions)	
	Deterministic	Uncertainty	Deterministic (NPV)	Uncertainty (ENPV)
1	$C_d=50$	$C_u=25$	$V_d=1.75$	$V_u=0.87$
0.95	$C_d=100$	$C_u=75$	$V_d=21.51$	$V_u=14.27$
0.90	$C_d=175$	$C_u=125$	$V_d=51.75$	$V_u=37.18$
0.85	$C_d=200$	$C_u=175$	$V_d=84.56$	$V_u=61.18$

5.4 Step 3: Flexibility Analysis

Using concept generation techniques inspired from Cardin et al. [15], flexibility to expand capacity is recognized as a strategy to deal with uncertain demand growth. The idea is to build less capacity at the start – to avoid over commitment and over capacity, and to add capacity based upon demonstrated demand. Key to this strategy, of course, is that the original design should enable capacity expansion easily.

The analysis considers two kinds of capacity expansion. First, it looks at the benefits of building up capacity incrementally at the main site. Second, it considers the further advantage of moving additional modules in the field, close to the demand sites, as way of lowering transportation costs, and further exploiting the benefits from a modular approach to design and management.

5.4.1 Flexible Modular Design – No Move

Instead of building a fixed plant of optimal size as previously considered, this flexible strategy starts with a small initial module and expands as desired. The question when it would be good to expand is answered by the decision rule. The following decision rule was embedded in the simulation spreadsheet: IF "*the difference between the observed aggregate demand and current capacity at this site is higher than a threshold value*" THEN "*the capacity using the modular design capacity is expanded*" ELSE "*do nothing*". Using an exhaustive enumeration technique, it is found that the threshold value 80% offers a better system performance among other threshold values.

5.4.2 Flexible Modular Design – With Move

This flexible design strategy allows the designers to add capacity away from the main site, and to place it in the field nearer the demand sites. The analysis has to implement two additional decision rules to explore this flexibility, to address two important questions: when should the modular plant be built for the first time at distance, and where should it be built?

The decision rule regarding the capacity expansion was: IF "*demand at each demand site reaches a certain threshold value as a parameter of the decision rule*" THEN "*a modular production plant can be built at the demand site*" ELSE "*do nothing*". This threshold value was tuned by conducting another comprehensive enumeration. The results show that the threshold value of 200% offers more economic value as compared to others.

The decision rule used regarding the geographical location for capacity expansion was: IF "*distance between the main production site and each demand site exceeds the maximum preferred coverage range*" THEN "*a modular production facility can be moved into the demand site*" ELSE "*do nothing*". To build extra modular plants at demand sites, a capacity expansion is triggered based on the decision rule embedded at each geographical site: IF "*the difference between the observed demand and the current capacity (i.e. unmet demand) at the demand site reaches certain threshold value*" THEN "*extra modular capacity is deployed*" ELSE "*do nothing*". Using exhaustive enumeration, the decision rule is tuned and the best threshold value is set to 80%.

Table 3 shows the improvement in multi-criteria performance metrics because of flexibility, as compared to the optimum fixed design for both kinds of flexibility examined here.

Table 3 Improvement of multi-criteria performance metrics due to flexibility with no learning

Criteria	Value ($ millions)			Improvement (%)	
	Optimum fixed design	Modular	Modular with move	Modular	Modular with move
ENPV	14.53	19.27	19.81	32.65%	36.40%
$VaR_{10\%}$	2.96	4.23	3.59	42.92%	21.28%
$VaG_{90\%}$	20.46	33.63	38.88	64.36%	90.04%

5.4.3 Multi Attribute Decision-Making

The best design alternative can be chosen based on many criteria. Some common economic metrics in project evaluation under uncertainty are shown in Table 4. The results correspond to the optimum fixed design with the economies of scale 0.95 and the flexible designs (with and without move) in terms of different learning rates. The aim is to choose a design based on the highest value for ENPV (or mean NPV), P10 VaR and P90 VaG, and smaller values for semi-standard deviation of NPV distribution and initial CAPEX.

Table 4 Multi-criteria decision making table considering α=0.95, figures are in million dollars

α=0.95	On-shore LNG production system design												
	Optimum	Flex 1: Flexible-no move			Flex 2: Flexible-with move			Best design			Value of flexibility		
	Fixed	Learning rate			Learning rate			Learning rate			Learning rate		
Criteria	(75 tpd)	0%	10%	20%	0%	10%	20%	0%	10%	20%	0%	10%	20%
ENPV	14.53	19.27	36.77	49.92	19.81	37.23	53.97	Flex 2	Flex 2	Flex 2	5.29	22.70	39.44
VaR	2.96	4.23	10.26	15.01	3.59	9.29	16.79	Flex 1	Flex 1	Flex 2	1.27	7.29	13.83
VaG	20.46	33.63	62.57	85.30	38.88	70.44	97.12	Flex 2	Flex 2	Flex 2	18.42	49.98	76.66
SSTD	10.85	2.54	1.91	1.61	2.76	2.54	1.63	Flex 1	Flex 1	Flex 1	8.31	8.94	9.24
Capex	60.44	27.50	27.50	27.50	27.50	27.50	27.50	Flex	Flex	Flex	N/A	N/A	N/A

5.5 Step 4: Sensitivity Analysis

This section investigates the sensitivity of the flexibility analysis to different economies of scale and learning rates. Table 5 shows the results: a) when economies of scale are stronger (α is smaller), the value of flexibility decreases. The reason is that strong economies of scale negate the value of deferring investments in capacity; b) when learning is greater, modules are cheaper, and flexibility is more valuable.

Table 5 Sensitivity of value of flexibility to different □ and LR

Economies of scale (□)	Flexible 1: modular design - no move			Flexible 2: modular design – with move		
	Learning rate			Learning rate		
	0%	10%	20%	0%	10%	20%
α=1	18.40	35.90	49.06	18.94	36.36	53.10
α=0.95	4.74	22.24	35.40	5.29	22.70	39.44
α=0.90	0.00	0.01	13.17	0.00	0.47	17.22
α=0.85	0.00	0.00	0.00	0.00	0.00	0.00

Overall, the value of flexibility depends mostly on four factors: a) uncertainty – the greater the uncertainty, the greater the value of flexibility, b) discount rate – which motivates the deferral of investment so as to minimize the present value of costs, c) economies of scale – which provide the incentive to build single big facilities at once, rather than smaller facilities developed in phases, leading to decreased value of flexibility, and d) learning effects that counterbalance economies of scale, in that they reduce the cost of implementing second and later addition of modules, and thus lead to improve the value of flexibility.

6 Conclusion

This study illustrates the value of flexibility in the design of production facilities under explicit considerations of uncertainty. It motivates the use of flexibility in engineering design as a paradigm to deal with uncertainty affecting lifecycle performance of engineering systems. The study represents an argument for a shift in the design paradigm away from the frequent focus on economies of scale focusing on the development and deployment of unitary large facilities that embody this advantage.

The paper relies on a structured four-step methodology inspired from existing literature [3]. It demonstrates the economic value of flexibility in the long-term design and deployment of production facilities subject to demand growth uncertainty. It considers the combined effects of economies of scale, learning, and the time value of money to highlight the economic benefits stemming from explicit considerations of uncertainty and flexibility. The case study concerns the prospects for LNG facilities in a target transportation market. The concepts are general, however, and can be applied to other distributed engineering systems sharing similar characteristics.

The results support the view that a flexible modular design can enhance economic performance compared to an optimum fixed design strategy. Furthermore, the flexibility to locate additional capacity beyond the main facility can further enhance the value of the system. Consideration of flexibility, however, adds another layer of complexity to the analytical problem. While an exhaustive search for the optimal design variables and decision rules is feasible here, considerations of more uncertainty sources, flexibility strategies (e.g. site abandonment, investment deferral), and more sophisticated decision rules can turn a tractable problem into a highly complex computational one. More work is under way to address these issues by combining meta-modeling and simulation-based optimization budgeting with stochastic programming techniques.

Acknowledgments. This work was supported by the National University of Singapore (NUS) Faculty Research Committee via MoE AcRF Tier 1 grants WBS R-266-000-061-133, and by the Agency for Science, Technology and Research (A*STAR) via the Singapore International Graduate Award (SINGA). The authors thank Wen Sin Chong, Ravindu Atapattu, and Dr. Kok Seng Foo from Keppel Offshore and Marine Technology Centre (KOMTech) for their contributions.

References

1. Kumar, S., Kwon, H.-T., Choi, K.-H., Hyun Cho, J., Lim, W., Moon, I.: Current status and future projections of LNG demand and supplies: A global prospective. Energy Policy 39(7), 4097–4104 (2011)
2. ECS. Western Australia Natural gas Demand and Supply - A Forecast. Economics Consulting Services (ECS) (2011), `http://www.latentpet.com/_content/documents/575.pdf`
3. de Neufville, R., Scholtes, S.: Flexibility in Engineering Design. MIT Press, Cambridge (2011)
4. Cardin, M.-A.: Enabling Flexibility in Engineering Systems: A Taxonomy of Procedures and a Design Framework. ASME Journal of Mechanical Design 136(1) (2014), doi:10.1115/1.4025704
5. Trigeorgis, L.: Real options: managerial flexibility and strategy in resource allocation. MIT Press (1996)
6. Wang, T.: Real Options in Projects and Systems Design – Identification of Options and Solutions for Path Dependency. Doctoral Dissertation in Engineering Systems, Massachusetts Institute of Technology, Cambridge, MA, United States (2005)
7. Fitzgerald, M.E., Ross, A.M., Rhodes, D.H.: Assessing uncertain benefits: a valuation approach for strategic changeability (VASC). In: INCOSE International Symposium. Citeseer (2012)
8. Ross, A.M.: Managing unarticulated value: changeability in multi-attribute tradespace exploration. Massachusetts Institute of Technology, Cambridge (2006)
9. Ross, A.M., Rhodes, D.H.: Using natural value-centric time scales for conceptualizing system timelines through epoch-era analysis (2008)
10. Özelkan, E.C., D'Ambrosio, A., Teng, S.G.: Optimizing liquefied natural gas terminal design for effective supply-chain operations. International Journal of Production Economics 111(2), 529–542 (2008), doi:`http://dx.doi.org/10.1016/j.ijpe.2007.02.028`
11. Grønhaug, R., Christiansen, M.: Supply chain optimization for the liquefied natural gas business. In: Innovations in Distribution Logistics, pp. 195–218. Springer (2009)

12. Andersson, H., Christiansen, M., Fagerholt, K.: Transportation planning and inventory management in the LNG supply chain. In: Energy, Natural Resources and Environmental Economics, pp. 427–439. Springer (2010)
13. Cardin, M.-A., Ranjbar Bourani, M., de Neufville, R., Deng, Y., Chong, W.S., Atapattu, R., Sheng, X.X., Foo, K.S.: Quantifying the Value of Flexibility in Oil and Gas Projects: A Case Study of Centralized vs. Decentralized LNG Production. Keppel Offshore and Marine Technology Review, Singapore (2013)
14. Savage, S.L.: The Flaw of Averages: Why We Underestimate Risk in the Face of Uncertainty. John Wiley and Sons (2009)
15. Cardin, M.-A., Kolfschoten, G.L., Frey, D.D., Neufville, R., Weck, O.L., Geltner, D.M.: Empirical Evaluation of Procedures to Generate Flexibility in Engineering Systems and Improve Lifecycle Performance. Research in Engineering Design 24(3), 277–295 (2013), doi:10.1007/s00163-012-0145

Optimising Complex Networked Systems Availability

Siew Yen Cheng and Yong Huang Seah

Abstract. With the Singapore Armed Forces (SAF) building up network-centric warfare capabilities, the reliance on the multiplier effects from inter-connectivity and collaborative operations among forces becomes increasingly critical. Availability has been applied typically at the system level, as a means to analyse the readiness and logistics effectiveness of the fighting force. Such a standalone system level Measure of Effectiveness (MOE) is no longer adequate to capture complex interdependency and ensure the readiness of networked systems in a holistic manner.

Today, the Optimised Decisions in Networks (ODIN) tool equips DSTA and the SAF with the ability to quantify networked system architecture and provide the means to identify critical links/bottle-necks that enhance design decision of the architecture. It provides us the means to examine network robustness and survivability under complex threat environment. ODIN seeks to perform resource (spares, manpower, equipment) optimisation at the network or System-of-Systems (SoS) level to ensure they are considered holistically to meet stringent demands. This paper aims to describe the methodologies and capabilities of ODIN. Such Systems Engineering approach could be similarly applied to the design of our smart cities to provide resiliency in design and best allocation of resource to meet the inter-dependencies and high degree of connectivity needed for utilities, transport and communications of today cities.

1 Introduction

Today modern cities exhibit high level of sophisticated living standards. This requires holistic and integrated urban planning across multiple agencies like utilities, transportation, communication and supporting infrastructure. Basics of water, gas and electricity supply network, efficient public road and rail transportation system,

---

Siew Yen Cheng · Yong Huang Seah
Defence Science Technology Agency, 1 Depot Road, 109679, Singapore

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_7

seamless network communication are just some of the essentials of today's smart cities. Ability to achieve such high level of integration requires a top level System-of-Systems (SoS) approach in planning and handling of inter-dependencies amongst the various supporting infrastructure networks to optimise resource allocation, especially in manpower and land scarce Singapore.

In similar fashion, the need for defence systems to be integrated into a System-of-Systems and operate seamlessly across functional and resource constraints is no less demanding. Our next generation Singapore Armed Forces (SAF) is seeing revolutionary changes in operation tempo, mission definition and combat service support. Systems are becoming more inter-connected and interdependent to leverage the network and information as force multipliers. For instance, a typical defence capability will consist of not one but several systems, made up of weapons systems connected to communication and sensor systems, with each possibly taking the form of a complex network. Planning done at the system (platform) level is no longer adequate to ensure mission success for such network centric operations.

The Optimised Decisions in Networks (ODIN) tool was developed to transform planning to support challenges in network centric operations. ODIN was developed with the ability to simulate complex network topology while incorporating the network systems operational profile, logistics maintenance support concept, system reliability and combat damage modeling. This tool aims to quantify inter-dependency and inter-connectivity across component systems in a networked system. With the means to quantify, ODIN enables one to identify weak links and optimise resource at a network level that ensures mission success for network centric operations.

The Systems Engineering approach and resource optimisation tool described in the paper was developed for the SAF, but is equally applicable and relevant for urban planning of today's smart cities.

2 System Level Availability

System level availability (Ao) is defined as the average availability of the system out in the field. Take the example of a sensor. Factors that influence the sensor system Ao covers not just the system inputs in terms of reliability and maintainability, but also the operational concept and logistics factors such as maintenance support concept, spares, technician quantity and available maintenance window (see Figure 1). Improving system availability involves not just optimal allocation of spare parts across maintenance agencies but it is also highly dependent on how in particular the sensor is being operated, its inherent system performance and the corresponding supporting maintenance factors.

The above factors can be translated into a quantifiable steady state Ao formula for a system of multiple LRUs indexed by k, k=1,…, k, as shown in the equation below. A Ao of 80% indicates that it is ready for mission on an average of 80 out of 100 hours. Statistically, it can also be interpreted as having an average of 8 out of 10 systems available.

$$\mathrm{Ao} = \frac{\mathrm{MTBD}}{\mathrm{MTBD} + \mathrm{MTTR} + \mathrm{WT}} = \frac{1}{1 + \dfrac{\mathrm{MTTR}}{\mathrm{MTBD}} + \mathrm{EBO}} \qquad (1)$$

Where MTBD is the mean time between demand and waiting time, $\mathrm{WT} = \mathrm{EBO} * \mathrm{MTBD}$ is defined by Little's Law.

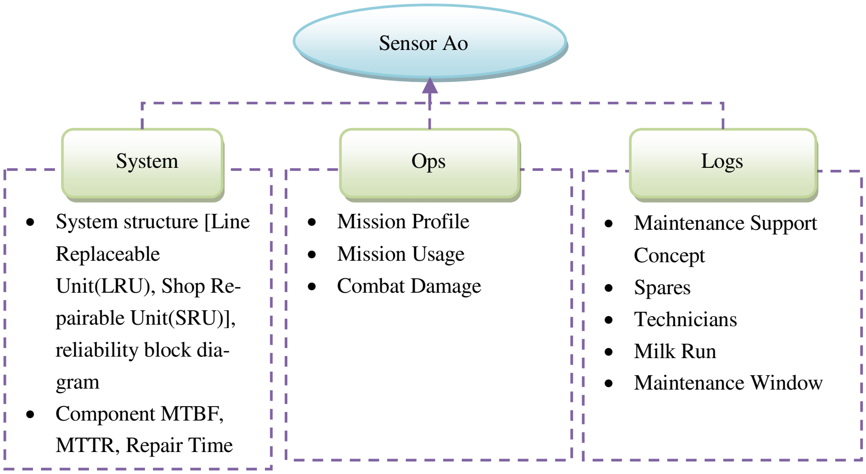


Fig. 1 Factors influencing the sensor availability

A modified version of the formula is proposed as follows (Lau, Song, See and Cheng, 2006):

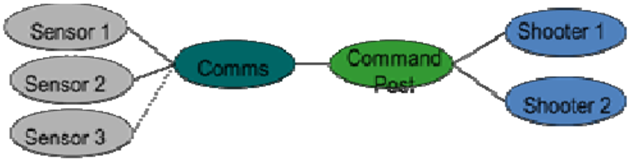
$$\mathrm{Ao} = \frac{1}{1 + \sum_{k} \left(\dfrac{\mathrm{EBO}_k}{\mathrm{N}_{\mathrm{sys}}} + \dfrac{\mathrm{UR}}{\mathrm{MTBF}_k} * \mathrm{QPM}_k * \mathrm{MTTR}_k \right)} \qquad (2)$$

Where EBO_k : Expected backorder of LRU_k;
MTBF_k: Mean Time Between Failures of LRU_k;
MTTR_k: Mean Time to remove and replace LRU_k;
$\mathrm{N}_{\mathrm{sys}}$: Number of system deployed
UR: Utilisation rate of system;
QPM_k: Quantity of LRU_k that the system has;

System level Ao is computed based on the summation across all the LRUs within the system. However, at a networked system level where the end-to-end mission requires collection of systems operating with certain interoperability and interdependency, such standalone system level Ao is no longer adequate.

3 Networked System Availabity

Networked system availability is defined as the availability of the interconnected systems at an end-to-end level. It quantifies the availability of having a link from one point to the other while having to route through the various component systems. Each of the component systems has its individual Ao defined by system level dependency on system, operational and logistics factors shown in Figure 1. Many often argue that such networked Ao can be obtained by simply multiplying them together using analytical formulae. This will derive a quick answer to the simple series-parallel type of networked system shown in Figure 2. However, such a method is very restrictive. First, typical networked systems are often meshed to meet the network redundancy requirements, and it is difficult to formulate the analytical equation. Second, it is not possible to capture the interoperability and interdependency that occur simultaneously across the multiple system types. The largest drawback lies in the analytical formulae multiplying the average of each component system Ao and hence losing the interdependency effect across systems that is the critical basis to the availability of a networked system. In the next section, the limitation of applying system level availability to an increasingly networked system environment is further illustrated.



Ao (Networked System) =

$$[1-(1-Ao_{\text{Sensor 1}}) * (1-Ao_{\text{Sensor 2}}) * (1-Ao_{\text{Sensor 3}})] * Ao_{\text{Comms}} * Ao_{\text{CommandPost}} * [1-(1-Ao_{\text{Shooter 1}}) * (1-Ao_{\text{Shooter 2}})]$$

Where Ao (Networked System): Availability of interconnected system end-to-end;

$Ao_{\text{Sensor i}}$: Availability of sensor system i;

Ao_{Comms}: Availability of comms system;

$Ao_{\text{CommandPost}}$: Availability of Command Post system;

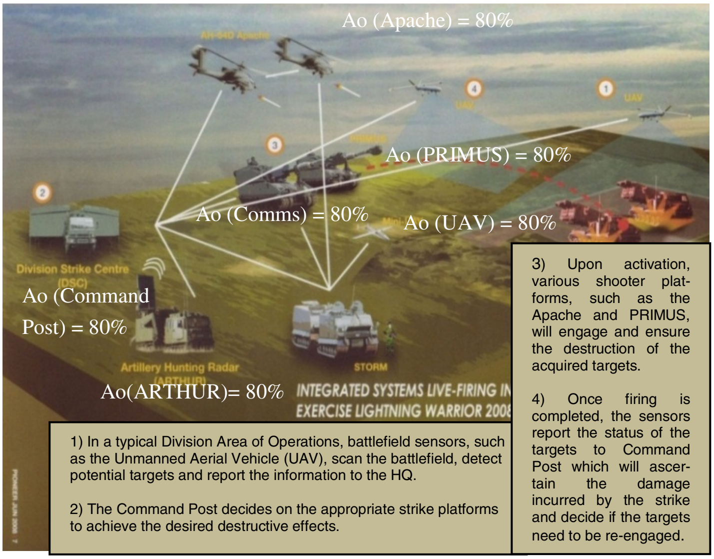
$Ao_{\text{Shooter j}}$: Availability of Shooter system j

Fig. 2 Analytical computation for simple networked system availability

3.1 Analysing Networked System Availability

Using an integrated system live firing exercise as illustrated in Figure 3, sensors in the form of a Unmanned Aerial Vehicle (UAV) or Artillery Hunting Radar (ARTHUR) are used to conduct battlefield surveillance and detect potential targets. Images of the ground surveillance are sent back to the command post via a communication network that allows the commander to decide on the appropriate strike platforms to take out the adversaries. From the command post, the target positions and information are sent via a communication network to the strike platform, which will engage and ensure the destruction of the acquired targets. It is evident that the mission success of acquiring and destroying the adversary is dependent on the simultaneous working of all systems types inclusive of communications networks. Should any of the systems be down, the mission will fail.

Typically, Ao, spares and resource are evaluated and allocated for each individual system; for example, Ao of 80% for each of the sensor and shooter systems. Such measurement is unable to reflect the interdependency of the various systems across the communication network for the mission. It may also potentially lead to under or over provision of resource and impact the logistical readiness of the systems.



1) In a typical Division Area of Operations, battlefield sensors, such as the Unmanned Aerial Vehicle (UAV), scan the battlefield, detect potential targets and report the information to the HQ.

2) The Command Post decides on the appropriate strike platforms to achieve the desired destructive effects.

3) Upon activation, various shooter platforms, such as the Apache and PRIMUS, will engage and ensure the destruction of the acquired targets.

4) Once firing is completed, the sensors report the status of the targets to Command Post which will ascertain the damage incurred by the strike and decide if the targets need to be re-engaged.

Fig. 3 Integrated systems live-firing exercise

A simplified "acquire & destroy" mission calculation is shown at Figure 4, where ARTHUR is used as the only sensor, PRIMUS (Self Propelled Artillery Gun System) as the weapon system, command post as the command and control centre, and a fixed communication network as the means of information and data transmission. Adopting a standalone system as the criteria for resource or maintenance support planning, the planner would ensure a Ao of about 80% for each of the individual system. However, from the "acquire & destroy" mission definition, it would require all the systems to be functioning together. If the planner's resource planning for each system is at 80% Ao, by simple multiplication the entire networked system is having maximum logistic readiness of only 40%. Therefore resource planning should be carried out at the networked system level. Planners can no longer perform their resource and maintenance support planning by treating each system as a standalone system. With the interdependency among the systems, the Ao of each system may no longer be treated independent of one another. Measurement of the performance of the networked system "acquire & destroy" mission needs to be performed within the model itself.

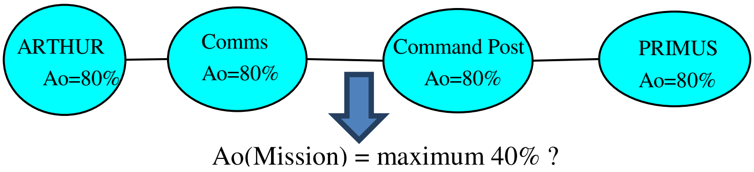


Fig. 4 Simple illustration of Integrated mission Ao computation

4 Extending Network Ao Computation

With such a complex network structure, system level Ao measurement can no longer suffice as a good Measure of Effectiveness (MOE) as it becomes more dynamic and largely dependent on the context. Two MOEs will first be defined before moving on to illustrate how these MOEs are used.

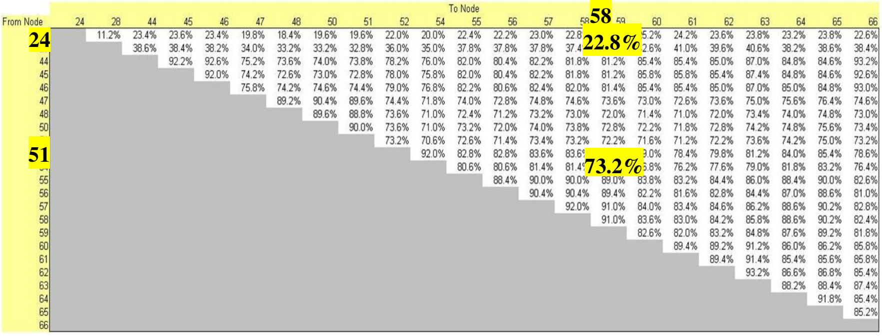
4.1 Mission Ao / Probability of Mission Ao

Mission Ao will see tighter integration between the operational and logistical contexts. This MOE requires the operational context to define how the operators had intended to interoperate the systems to ensure mission readiness. This mission Ao is highly dynamic and dependent on mission definition.

As illustrated in Figure 4, the mission Ao is defined as the "acquire & destroy" mission. It measures the probability of having sensors acquire the targets and transmitting the information to the appropriate shooters for them to take out the adversaries simultaneously. Mission Ao can also take the form of division to brigade Ao which measures the end-to-end availability from division to brigade by factoring the means for commander to communicate to ensure mission success.

4.2 Matrix of System-to-System Ao

For a large communication network, single networked system Ao is not representative. Instead, there is a paradigm shift towards the use of upper triangular matrix of multiple source-sink pairs as illustrated in Figure 5. This MOE allows one to evaluate each pair of system-to-system Ao to identify the weak links and bottlenecks at a glance; for example, system-to-system availability of 24 to 58 is low at 22.8% while, system-to-system availability from 51 to 58 is at 73.2%. In addition, system 24 is observed to have low availability to any other systems (first row of matrix), pinpointing to system 24 as one of the key bottleneck.



											To Node											
From Node	24	28	44	45	46	47	48	50	51	52	54	55	56	57	58	59	60	61	62	63	64	65	66
24		11.2%	23.4%	23.6%	23.4%	19.8%	18.4%	19.6%	19.6%	22.0%	20.0%	22.4%	22.2%	23.0%	22.8%	[illegible]	[illegible]	24.2%	23.6%	23.8%	23.2%	23.8%	22.6%
[illegible]			38.6%	38.4%	38.2%	34.0%	33.2%	33.2%	32.8%	36.0%	35.0%	37.8%	37.8%	37.8%	37.4%	[illegible]	[illegible]	41.0%	39.6%	40.6%	38.2%	38.6%	38.4%
44				92.2%	92.6%	75.2%	73.6%	74.0%	73.8%	78.2%	76.0%	82.0%	80.4%	82.2%	81.8%	81.2%	85.4%	85.4%	85.0%	87.0%	84.8%	84.6%	93.2%
45					92.0%	74.2%	72.6%	73.0%	72.8%	78.0%	75.8%	82.0%	80.4%	82.2%	81.8%	81.2%	85.8%	85.8%	85.4%	87.4%	84.8%	84.6%	92.6%
46						75.8%	74.2%	74.6%	74.4%	79.0%	76.8%	82.2%	80.6%	82.4%	82.0%	81.4%	85.4%	85.4%	85.0%	87.0%	85.0%	84.8%	93.0%
47							89.2%	90.4%	89.6%	74.4%	71.8%	74.0%	72.8%	74.8%	74.6%	73.6%	73.0%	72.6%	73.6%	75.0%	75.6%	76.4%	74.6%
48								89.6%	88.8%	73.6%	71.0%	72.4%	71.2%	73.2%	73.0%	72.0%	71.4%	71.0%	72.0%	73.4%	74.0%	74.8%	73.0%
50									90.0%	73.6%	71.0%	73.2%	72.0%	74.0%	73.8%	72.8%	72.2%	71.8%	72.8%	74.2%	74.8%	75.6%	73.4%
51										73.2%	70.6%	72.6%	71.4%	73.4%	73.2%	72.2%	71.6%	71.2%	72.2%	73.6%	74.2%	75.0%	73.2%
[illegible]											92.0%	82.8%	82.8%	83.6%	83.6%	[illegible]	[illegible]	78.4%	79.8%	81.2%	84.0%	85.4%	78.6%
[illegible]												80.6%	80.6%	81.4%	81.4%	[illegible]	[illegible]	76.2%	77.6%	79.0%	81.8%	83.2%	76.4%
55													88.4%	90.0%	90.0%	89.0%	83.8%	83.2%	84.4%	86.0%	88.4%	90.0%	82.6%
56														90.4%	90.4%	89.4%	82.2%	81.6%	82.8%	84.4%	87.0%	88.6%	81.0%
57															92.0%	91.0%	84.0%	83.4%	84.6%	86.2%	88.6%	90.2%	82.8%
58																91.0%	83.6%	83.0%	84.2%	85.8%	88.6%	90.2%	82.4%
59																	82.6%	82.0%	83.2%	84.8%	87.6%	89.2%	81.8%
60																		89.4%	89.2%	91.2%	86.0%	86.2%	85.8%
61																			89.4%	91.4%	85.4%	85.6%	85.8%
62																				93.2%	86.6%	86.8%	85.4%
63																					88.2%	88.4%	87.4%
64																						91.8%	85.4%
65																							85.2%
66																							

Fig. 5 Illustration of the upper triangular matrix tabulation of System-to-System Ao

5 Introducing ODIN

Recognizing the limitations and needs, ODIN was developed to model and evaluate the Ao at a networked system level to reflect the inter-connectivity among systems. Critical links and bottlenecks can hence be easily identified by measuring end-to-end Ao between system nodes. The tool also provides the ability to evaluate the survivability of the network when subjected to combat damage scenarios. In addition, ODIN allows the evaluation of network robustness when subjected to various what-if scenarios; for instance, is the network still able to fulfil its original mission intent when one or two nodes are down?

ODIN provides the capability to optimise the spares and resource at a network level. In this way, resource are no longer allocated uniformly across all systems. Instead, systems found to be the bottleneck or weak links identified from end-to-end evaluation will be allocated higher resource level to better optimise network availability in a more holistic manner. This trade-off ensures that mission readiness can be achieved at the end-to-end level using the most cost effective approach.

5.1 ODIN Methodologies

There have been extensive publications on the performance of network analysis or quantifying resiliency in the context of networks. Research on network reliability quantifies the probability that the network performs its intended function for a specific mission time under known, normal operating conditions (Elsayed, 1996). Other approaches look at quantifying the resilience of the network when subjected to external causes of component failure such as potential catastrophic failures due to attacks, disasters etc (Whitson and Ramirez-Marquez, 2009). Within the ODIN model, measurement of network performance with reliability and probabilistic combat damage (external factors) as failure sources was adopted. Dependencies of the various systems are also looked into to aid in identifying any correlation and weak links.

In a networked system domain, the interaction of the various systems can be viewed as a network with multiple nodes and links. In ODIN, the source node is defined as the origin and the sink node as the destination. For the network, performance is measured in terms of the ability to pass through from the source to the sink without any interruption from any broken links or nodes. Each system is represented as a node and it can be mapped from the source to the sink system with its network Ao measured. Each individual system availability is affected by its inherent system/component reliability, external probabilistic combat damage together with its unique logistical factors (e.g. maintenance support concept, finite resources) while operating together under the networked system concept.

The mission success of a networked system often requires some degree of interoperability among the individual systems which may be physically sited in different network layers. In order to address this concept of operations, multi-layered networks are modelled within the same model through the concept of network mesh and sub-layers. For example, in the "acquire & destroy" mission, by modelling the communication network as a common mesh layer, the interdependency can be modelled without complicating the network and yet achieve the effect of system dependency on the communication network.

With the use of such a network mesh, users are able to model the interaction and interdependency on different network layers within the same model. The run time of the model is significantly reduced as the network mesh provides the means of decomposition of large complex network layers. Each layer can be computed independently and their interdependency merged rather than computing for a huge complex network (see Figure 6). In addition, rule sets that govern unique routing of each of the network layers can be customised and implemented to capture the network performance accurately.

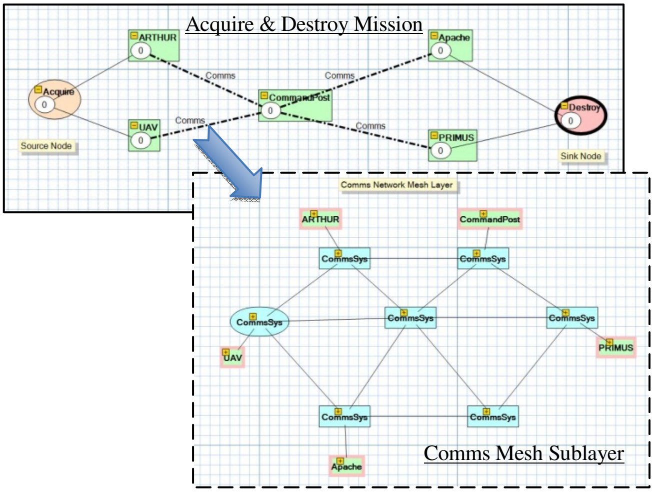


Fig. 6 Illustration of the use of network mesh for modelling the interdependency of different network layers

5.2 *ODIN Tool Architecture*

ODIN is made up of numerous libraries and engines, each designed with unique methods and functionalities. Three main engines, namely the process simulation engine, resource optimisation engine and network search computation engine, are integrated to provide the full capabilities of ODIN. The process simulation engine adopts a Monte Carlo simulation method to evaluate the dynamics and stochastic system failures, logistics supply chain, repair process and mission profile. The resource optimisation engine built using mathematical algorithms provides optimum spares recommendation. Finally, the network search computation engine adopts path searching techniques to compute and evaluate the interdependency across networked systems. Together, they provide a holistic solution to the optimisation of complex networked system availability. A high level architectural view of ODIN is illustrated in Figure 7.

Typically there are two types of failure demand: one arising from component reliability failures and the other due to combat damage as a result of executing a combat mission. The unique nature of each problem type uses different methods in providing a solution in ODIN. Failures in reliability are approximated with mathematical formulation while combat damage failures are approximated via stochastic. For reliability spares optimisation, mathematical formulation - an approximate to Palm's theorem and classical Multi-Echelon Inventory theory (Lau et al., 2006; Sherbrooke 1992; Alfredson, 1997) - is used. The combination of metaheuristic algorithm with Monte Carlo simulation (Dubi, 2000) is adopted for combat damage spares optimisation.

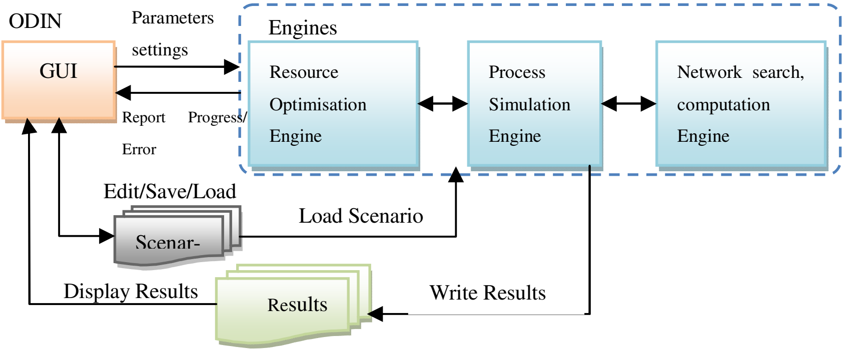


Fig. 7 An Architecture View of ODIN

6 Implementation and Case Studies

6.1 *Case Study 1: Networked System Architecture Evaluation*

As part of the architectural evaluation of the robustness of the Networked Air Defence design (see Figure 8) in meeting its mission objectives, end-to-end network availability from sensor to C2 (command and control) to shooters was performed. Several key considerations were factored in; for example, sensor network since there was no dedicated sensor-to-shooter pairs. In addition, shooters were dispersed across large geographical locations and linked back to the central C2 system. Moreover, there was the need to handle the IT infrastructure and communication equipments to provide the connectivity among sensor, shooter and C2. Adding to the complication was the different network configurations across different mission phases. All these were modelled through ODIN where multiple network layers were inter-connected and inter-linked to provide end-to-end mission readiness.

Due to the different capabilities of the sensors and shooter in terms of range and threat types, no single mission Ao could be defined. Instead, a matrix of MOE based on the threat and campaign type was used. For example, against threat X, availability was measured from sensor A or B to Shooter I or II. ODIN enabled the mission readiness of Networked Air Defence to be evaluated in totality despite the independent management of individual systems. This ensured robustness in networked system architecture design with respect to connectivity between the component systems. This was achieved through the quantification and identification of weak links and/or vulnerabilities which enabled the optimisation of the Networked Air Defence Ao through improved connectivity configuration and incorporation of system redundancy.

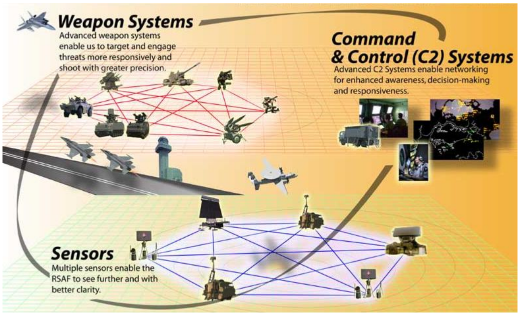


Fig. 8 Evaluation of robustness of Networked Air Defence architecture

6.2 Case Study 2: Networked System Resource Optimisation

A C4 (command, control, communications, and computer systems) system consists of many component systems connected together in a functional relationship. Typically, Ao is measured and resource are catered for at a system or node level. However, it does not provide a commander with a sense of the mission readiness. Hence, this study aimed to evaluate end-to-end network Ao from division to brigade level by piecing together the radios, phones, Command Control Information System (CCIS) to trunk communications equipment. ODIN provided the means to quantify the network Ao down to data versus voice. Such an approach ensured that the spares deployment from different equipments were well balanced with respect to end-to-end availability.

Optimising end-to-end Ao requires tradeoff across multiple factors. For network architecture, it involves deciding between the number of radio links versus the number of radio redundancies available for each system node. There is also tradeoff among the various system configurations as well as the logistics input of spares deployment and support to determine the response to system and network downtime.

With ODIN, the modelling approach takes a step back to look at the fundamental functional level. Instead of the physical series-parallel reliability block diagram modelling, functional routing within and across the systems are modelled so that the system configuration design and differentiation between the voice (V) and data (D) routes can be accurately captured. Figure 9 shows the different possible routing paths to reach end-to-end between the voice (V) to voice (V) and data (D) to data (D) system nodes.

Through such detailed modelling, overall end-to-end network Ao can be optimised globally across various factors including increased client redundancy, improved response time in spares support, review of system configuration design to achieve spares optimisation across systems, as well as operations and logistics at the network level. It involves the levelling of resources across the different component system nodes such as providing identified bottlenecks with higher resources. Results have demonstrated a 10% improvement in overall network Ao. Most importantly, it removes the previous siloed approach which is not only tedious and computationally hard to analyse across the many interacting factors, to today a top level global approach in an automated, elegant and exact solutioning.

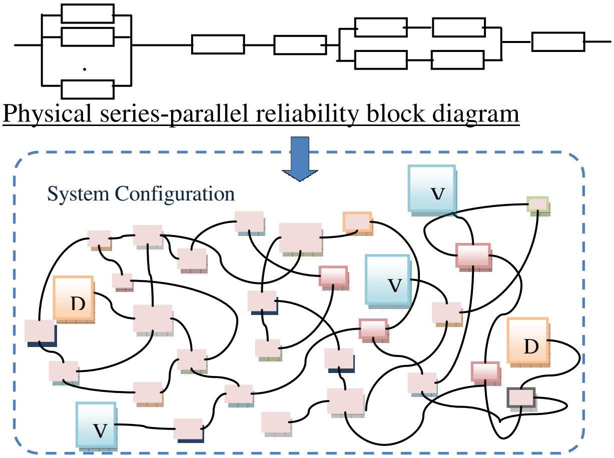


Fig. 9 Illustration of physical series-parallel reliability block diagram modelling versus functional network routes modelling

7 Potential Applications

Through quantifying the availability of the network, identifying weak links and managing system dependencies, ODIN can be used for the following potential applications:

1. Front-end planning tool in the design of resilient networked system architecture or for network planning;
2. Multi-resource optimisation at networked system level to obtain cost effective solutions while ensuring end-to-end mission readiness.
3. Logistic planning tool to aid the commander in verifying that the logistics plans are able to support operational plans and vice versa.

8 Conclusion

ODIN is a Systems Engineering tool that DSTA developed to support own work on complex networked capability design and realization. It equips DSTA with the ability to evaluate end-to-end availability of networked system architecture and captures the inter-connectivity and interdependency across the various systems. It allows one to identify the vulnerabilities and resilience of architecture towards threats. Most importantly, spares and resource optimisation can now be done at a networked system level that results in cost effective solutions to ensure end-to-end mission readiness. Such Systems Engineering approach could be similarly applied for the design of our smart cities. This provide resiliency in design and best allocation of resource to meet the high degree of connectivity and inter-dependencies needs for utilities, transport and communications of today cities.

Acknowledgements. The authors would like to acknowledge and credit Mr Soh Zhao Sheng, Mr Adrian Chow Meng Hwee and Ms Ong Xiu Hui from the DSTA ODIN development team together with their users in Singapore Army - Headquarters Maintenance and Engineering Support (HQMES) for their contribution to the tool’s success.

References

1. Alfredson, P.: Optimization of multi-echelon repairable item inventory systems with simultaneous location of repair facilities. European Journal of Operational Research 99(3), 584–595 (1997)
2. Dubi, A.: Monte Carlo applications in systems engineering. John Wiley & Sons Inc., New Jersey (2000)
3. Elsayed, A.E.: Reliability engineering. Addison Wesley Longman, Boston (1996)
4. Lau, H.C., Song, H.W., See, C.T., Cheng, S.Y.: Evaluation of time-varying availability in multi-echelon spare parts systems with passivation. European Journal of Operational Research 170(1), 91–105 (2006)
5. Sherbrooke, C.C.: Optimal inventory modelling of system: multi-echelon techniques. John Wiley & Sons Inc., New Jersey (1992)
6. Whitson, J.C., Ramirez-Marquez, J.E.: Resiliency as important measure in network reliability. Reliability Engineering and System Safety 94(10), 1685–1693 (2009)
7. Cheng, S.Y., Seah, Y.H.L.: Optimising complex networked systems availability. DSTA Horizons. Defence Science and Technology Agency, Singapore (2014)

Engineering Singapore's Land Transport System

Eng Seng Chia

Abstract. There is a wide array of public transportation in Singapore. Travelling from one part of Singapore to another is a breeze thanks to a highly penetrative public transport system. The three main modes of public transport in Singapore are the Mass Rapid Transit (MRT), buses and taxis. Buses are, by far, the most common form of public transportation in Singapore, followed by the MRT and taxis respectively. Public buses serve almost every part of Singapore, making it the most extensive form of public transportation, while the MRT provides speed and efficiency, especially during peak hours.

This case examines the public transportation of Singapore from a large scale system engineering perspective and aim to help the reader to understand the complexities of such a system: its mission, vision and goals, boundaries, and promoting public transport through a systems approach.

1 Formulation of Goals

In any large scale system, the formulation of goals is very important for various reasons: goals give the system a target to aim for; goals help the system to allocate resources and concentrate its time and effort; goals provide the motivation for the stakeholders to preserve; goals help the system to establish priorities. Very often, systems forget or ignore their goals, moving aimlessly in circles. Goals are usually derived from the system's mission and vision.

2 Mission

Mission or purpose is the reason for the system's existence. The mission statement help the organization make decisions that are in alignment with its core purpose.

Eng Seng Chia
College of Alice and Peter Tan,
National University of Singapore,
8 College Avenue East, B1-50, Singapore 138615
e-mail: `aaron_chia@nus.edu.sg`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_8

The Land Transport Authority (LTA) is a statutory board under the Ministry of Transport that spearheads land transport developments in Singapore. It plans the long-term transport needs of Singapore, taking care of those who drive as well as those who take public transport. LTA's mission is to "**Connecting people and places, enhancing travel experience**" (Land Transport Authority, 2014)

3 Vision

Vision is a clear image of your desired future. Vision answers the question, what is the final result we want? It is a picture of the future the system seeks to create. It gives shape and direction to the organization's future and helps people set goals and prioritize strategies for moving the organization closer to its desired results. Once one is clear on the vision, one can then focus on one's strategies that are in alignment with the results one wants to achieve. LTA's vision is "a people-centered land transport system." (Land Transport Authority, 2014).

4 Goals

From the mission and vision, the following goals/objectives are derived (Public Transport@SG, 2014):

a. To deliver a land transport network that is integrated, efficient, cost-effective and sustainable to meet the nation's needs.
b. To plan, develop and manage Singapore's land transport system to support a quality environment while making optimal use of our transport measures and safeguarding the well-being of the travelling public.
c. To develop and implement policies to encourage commuters to choose the most appropriate transportation mode.

LTA's strategy to provide an attractive public transport system includes (Land Transport Authority, 1996)

a. managing demand for private transport through ownership and usage restraint to ensure sustainable mobility into the long term,
b. integrating transport planning with land use planning for a sustainable outcome due to land scarce Singapore,
c. keeping roads safe and the air clean together with the Ministry of the Environment and Water Resources (MEWR)
d. reviewing vehicle standards regularly, public education and enforcement programmes together with the Traffic Police.

In the 2008's Land Transport Masterplan, the vision is to work towards a more people-centred land transport system that will meet the diverse needs of an inclusive, liveable and vibrant global city. Three key strategic thrusts were identified (Land Transport Authority, 2008):

a. Making public transport a choice mode
b. Managing road usage
c. Meeting the diverse needs of the people

We will see how these goals were translated into strategic thrusts and how they permeated everything that LTA does to achieve the system's goals. Last year, another Land Transport Master Plan 2013 was laid out to address the changing expectations of commuters and to achieve greater integration in common spaces.

5 Boundaries, Complexities and Integration

A boundary for a system is a periphery or a temporary limit at which some noticeable change in the application behaviour happens. It is a line or plane indicating the limit or extent of something. Some of these boundaries create conditions that are due to the system's intended behaviour or specifications, or business rules (e.g. "for getting a driver license, the applicant must be at least 18 years old"). Some are "implicitly" created e.g. importing foreign talents create new social and cultural boundaries. Hence it is very important to know what the boundaries and their conditions are in a large scale system so that the planner chooses how to deal with them. For example, in the past, one boundary condition was not to build housing in industrial estates due to pollution and possible health problems. This resulted in workers spending time transporting from their homes to their workplace. With better pollution control as well as building more gardens around industrial areas, it is possible to remove this boundary condition especially in a land scare country (which is another boundary) like Singapore.

Complexity arises in situations where an increasing number of independent variables begin interacting in interdependent and unpredictable ways. There are three types of complexity: system complexity (multiple connections and ways of implementing it); behavioural complexity (goals agreement and tradeoffs and multiple stakeholders) and; dynamic complexity (boundaries change with time). In many governmental projects such as transportation, the size and the many subsystems contribute to system complexity, the many stakeholders involved contribute to behavioural complexity, and the changing environment for such long time projects contribute to dynamic complexity.

Planning, systems engineering, systems thinking and dynamics, systems architecting are some of the ways to deal with complexity and boundaries. Some of these are described below.

6 System Complexity: Integration of Land Use, Town and Transport Planning

Proper land use planning improves accessibility to goods and services and hence reduces the need for transportation. In Singapore, the majority of commercial activities is planned to be located within the city centre to create a critical mass of activities and help position Singapore as a global business hub (Sapuan, 2007).

To allow people to live close to this concentration of employment opportunities, more housing is being introduced into and near the city centre. Outside the city, more employment areas are planned near suburban residential areas; for example, industrial estates and commercial hubs are situated at the fringe of housing estates to reduce commuting and lessen transport demand.

Together with other agencies, LTA integrates urban development with transport planning. Having a proper mix of developments and the highest building densities concentrated at and around MRT stations will ensure maximum accessibility, for commuters to key nodes of employment, housing, leisure and other social activities. Commuter facilities and building developments will be fully integrated. More high rise developments are being targeted near MRT stations. Radial lines will provide direct links to the city centre, while orbital lines will serve those travelling from one place to another outside the city centre. This will enhance accessibility and reduce travel times, making public transport more attractive and reducing reliance on private transport.

7 Dynamic Complexity: Integration over Time

In 1996, a White Paper on “A World Class Land Transport System” was conceived to guide Singapore’s land transport development (Land Transport Authority, 1996). However, the environment has changed as her population has grown and become more diverse. By 2020, she expects that travel demand would increase from the current 8.9 million journeys a day to about 14.3 million journeys a day (Land Transport Authority, 2008). The public transport mode share during the morning peak hours has declined from 67% in 1997 to 63% in 2004 and 59% in 2009 (Channel News Asia, 26 Oct 2009). Fortunately, it has risen to 63% in 2012 (LTA, 2013). Singapore also has to better meet the needs of her greying population, the higher expectations of her people as well as the less privileged and lower income group. Given these constraints and limited land of just over 700 square kilometers, public transport would be the most efficient means (Land Transport Authority, 2008). With this in mind, Land Transport Authority (LTA) embarked on a comprehensive Land Transport Review in October 2006 cumulating in the Land Transport Masterplan in 2008.

Hence some of the strategic thrusts from the 1996’s White paper need to be changed and the strategic thrusts in the 2008’s Land Transport Masterplan, reflects some of these changes. In 1996, the strategic thrusts were integrating transport and land use planning; expanding the road network and maximizing its capacity; managing demand of road usage and; providing quality public transport choices. In 2008, they were making public transport a choice mode; managing road usage and; meeting the diverse needs of the people. The need to manage the demand for road usage remains. But the latter emphases more on encouraging public transport and meeting different customers’ needs.

The environment in Singapore has changed significantly since 2008. Commuters expect to be more connected and travel faster and in comfort. More goods need to be transported. There will also be tighter land constraints (LTA, 2013 Land Transport Master Plan). The 2013's Land Transport Masterplan was conceived to meet these challenges. In essence, it strives to connect people to more places where they work, live and play, improve travel reliability, comfort and convenience and build and the transport system with the well-being of our diverse community at heart and give more consideration to how it enhances the common living space (LTA, 2013 Land Transport Master Plan).

8 Behavioural Complexity

Land transport is a matter that affects everyone. A people-centered land transport system must be planned with the community in mind. Apart from LTA and several other governmental organizations such as the Urban Redevelopment Agency, other stakeholders include transport operators, employers, commuters etc. To solicit and integrate their interests, LTA, in conceiving its masterplan, obtained contributions from a broad spectrum of people including students, workers, employers, commuters, transport operators, ordinary Singaporeans and experts; at home and abroad. In total, more than 4,500 people contributed their time, energies and ideas to the plan (Land Transport Authority, 2008). Going forward, the new Land Transport Community Partnership Division in LTA will have dedicated teams assigned to each constituency to engage the community more closely on the ground. LTA will also launch a Community Partnership Programme to invite grassroots leaders to discuss and share their views on transport policies and plans.

9 Behaviour Complexity: Integration across Organizations

One of the greatest challenges in large scale systems engineering is the integration of the different organizations. An understanding of organizations and their ecosystems would certainly be helpful to integrate them. It would certainly not prudent to merge all the entities involved in the system, as this will result in huge bureaucracies and hierarchies, not in the least having to resolve the politics involved. However, where interfaces and interactions between different entities remain high, it makes sense to combine these entities. Hence LTA was created in September 1995 to spearhead improvements to the land transport system through integrated planning, development and management of land transport policies and infrastructure. It was formed through the merger of four public sector entities, namely: Registry of Vehicles, Mass Rapid Transit Corporation, Roads & Transportation Division of the Public Works Department and Land Transport Division of the then Ministry of Communications. The formation of the LTA allows for better integration of the functions of planning, development, implementation and management of all transport infrastructure and policies. Such integration addressed the need for a more co-coordinated approach to planning and building Singapore's land transport system.

10 Sustainable Development

Achieving sustainability is one of the most important challenges facing society. Addressing complex issues related to sustainability requires systems integration of multiple disciplines including social, economic and environmental concerns. LTA recognizes the social, environmental and economic principles of sustainable development as a strategic objective.

A world class land transport system is very costly to build and operate. Hence a framework which serves as a social contract based on partnership between the Government, the operators and commuters was devised (Land Transport Authority, 1996). This framework is based on the principles that fares have to be realistic and revised periodically to account for justifiable cost increases to maintain acceptable service level, operating costs must be recovered to make it viable, and that a need for a sustainable policy on asset replacement. This means that with the government funding the assets, operators need not have to come up with a huge investment and be able to cover operating costs. They should also use some of the revenue to provide better service. To prevent public transportation from being unaffordable, the Public Transport Council scrutinize the fares and has to balance between the need to ensure that fares remain affordable to the public and operations are commercially viable in the long term. Figure 1 illustrates the economic/ social framework for a sustainable land transport system.

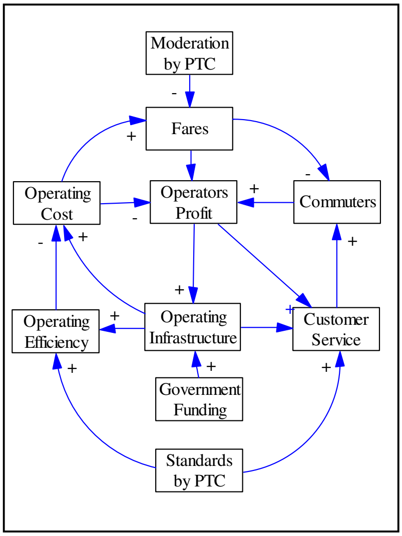


Fig. 1 Economic and Social Framework for a Sustainable Land Transport System

The basic bus service market is gradually being opened up to allow competition for the bus market i.e. competitive tendering for the right to operate a package of bus services. To encourage greater efficiency and service improvements for the benefit of commuters, greater contestability in the Rapid Transit System (RTS) industry is achieved by issuing shorter operating licences for future RTS lines, compared to the 30-year licenses today.

LTA works with the Ministry of the Environment (ENV) and the Traffic Police (TP) to keep the roads safe and the air clean. Regular reviews of vehicle standards, public education and enforcement programmes are held. In addition, noise alleviating measures such as underground roads and expressway noise barricades are also implemented.

LTA ensures that all vehicles in Singapore comply with rules, regulations and technical requirements to ensure that these are built to acceptable international standards and are safe for use on the roads. LTA also enforces exhaust emission standards for all vehicles by requiring regular inspections. The transport sector is the third largest consumer of energy after industry and building sectors and accounts for 15% of Singapore's total carbon dioxide emissions, with private cars as the chief contributor (35%) and public transportation (36%) which includes taxies, buses and MRT/ LRT (LTA, 2013 Household Interview Travel Survey 2012). It encourages energy efficiency and reduces carbon emissions by promoting the use of public transport and more energy efficient vehicles. Rebates are given to encourage the use of green vehicles by narrowing the cost difference between such vehicles and conventional vehicles. In addition, environmentally sustainable practices in the planning and development of transport infrastructure are adopted.

11 Increasing Capacity through Use of Technologies

LTA uses technology to maximize the capacity of our roads by upgrading and installing intelligent traffic management systems. It has installed the Green Link Determining (GLIDE) intelligent traffic light system to control all traffic signals along arterial roads. As traffic flow changes, it responds by adjusting the (traffic) green time accordingly. GLIDE also links traffic signals at neighboring junctions so that motorists can travel from one junction to another with minimal stops. "Virtual slip roads" are also implemented by allowing motorists to turn left at selected traffic light junctions even though the red signal is on, similar to the American right-turn-on-red system. Adopting traffic monitoring systems like the Automatic Network Travel Time System (ANTTS) and close circuit televisions to monitor traffic conditions in real-time. Junction Electronic Eyes System (J-Eyes) and the Expressway Monitoring Advisory System (EMAS) are also used to cover more junctions and monitor traffic conditions on arterial roads respectively. LTA further develops these systems to provide navigational and traffic information to commuters through the radio, telephone hotlines or variable message signs. These measures increase the capacity of roads and junctions and ensure smoother traffic flow.

Another system that LTA has introduced is the Electronic Road Pricing (ERP). ERP is an electronic system of road pricing based on a pay-as-you-use principle. It is designed to be a fair system as motorists are charged when they use the road during peak hours. Traffic conditions on the expressways and roads where the

ERP system is in operation are reviewed quarterly and its rates adjusted where necessary to minimize congestion on the roads.

12 Vehicle Growth

To keep traffic flowing smoothly, LTA adopted a holistic approach that includes road expansion as well as managing demand for road use by controlling vehicle growth and restraining usage. One such system of restraining usage through car ownership is the Vehicle Quota System (VQS). It is a system which allows the government to control the amount of cars on the road. The quota is reviewed on a regular basis and is set per month, based on the road conditions and amount of cars permanently taken off the road in that month (by scrapping, exporting or otherwise). The VQS is controlled by the Certificate of Entitlement, or COE. Each month, the LTA will determine how many vehicles are allowed on the road according to the VQS. The LTA will then issue that amount of Certificates of Entitlement (COE) and are sold through an open bidding system.

Parking policy is another lever that will restrain car usage by reducing parking supply in the city gradually over time and increasing parking charges. Parking Guidance Systems (PGS) are also implemented to guide motorists to the nearest building with available parking to reduce circulating traffic looking for parking lots.

13 Public Transport System

To make public transport a choice, the land transport system must be able to provide various types of public transport services, ensure that these services are customer oriented, and integrate them to ensure a seamless journey.

14 Choice of Public Transport Services

Providing an attractive public transport system is the cornerstone of Singapore's land transport strategy. One way is in providing a comprehensive range of public transport services, each being developed to the highest quality commensurate with the fares charged, and all well integrated to provide a seamless journey. The choices include:

a. Mass Rapid Transport (MRT) to serve heavy transit corridors;
b. Light Rail Transit (LRT) systems to serve as feeders to the MRT network;
c. Buses to continue serving the less heavy corridors to complement MRT-LRT network;
d. Premier bus services like BusPlus to provide higher grade of bus service; and
e. Taxis to provide car-like services.

15 Customer Oriented Services

LTA also works with the operators to make public transport more customer-oriented. Examples of such measures include (Land Transport Authority, 1996):

a. improving travel times through more bus priority schemes like priority at traffic light junctions and bus lanes. It is mandatory for motorists to give way so that buses can come out of bus bays without delay, and buses will enjoy signal priority over other vehicles at major junctions in the city.
b. improving commuter facilities by providing bus arrival times at bus stops, over the phone or Internet;
c. providing better bus services by refurbishing bus stops, providing more covered linkways from their flats to improve accessibility of the public transport system;
d. encouraging higher operational efficiency through the use of Global Positioning System (GPS) to pinpoint the location of buses to significantly improve scheduling.

With these measures, commuters can expect a speedier and smoother ride on the bus. The reliability of bus travel will also improve as buses enjoy greater right of way on the roads.

The MRT network which serves heavy traffic corridors, is also being expanded the target is to have a coverage as comprehensive as the London Tube or Paris Metro. Within the Central Area, a commuter will be able to access a RTS station within five minutes walk on average. Capacity on the existing lines will also be enhanced by increasing the frequency of trains. Commuters can look forward to better connectivity and a more comfortable ride on the trains (Land Transport Authority, 1996).

Taxis offer personalized service and supplement the high end of public transport. Measures to improve this service include periodic evaluation and publication of the performance of taxi operators and incentives for the best performing fleet operator; differential pricing to address the acute shortage of taxis during peak periods; Flexibility for operators to offer a wider variety of services; and using GPS technology to improve the radiophone service to better match demand and supply.

16 Integration of Public Transport Services

To enhance the integration and efficiency of public transport services, the LTA took on the role of a central bus network planner in 2009. It plans the public transport network from the commuters' perspective, focusing on the 'total journey' experience of the commuters. LTA enhances the hub-and-spoke system, so that the bus and rail services work in partnership. There will be more frequent and direct feeder bus services so that commuters can reach the transfer hubs quickly, and enjoy seamless and efficient transfers to the MRT or trunk buses to continue with their journeys. Moreover, more fully integrated transport hubs where bus interchanges and RTS stations are co-located with retail and commercial activities are being built. To facilitate transfers, a distance-based through-fare structure was adopted by 2009 so that commuters is charged a fare based on the total distance travelled in a journey, without incurring a transfer penalty when they switch between buses or between the bus and MRT. This will encourage commuters to take the most efficient routes.

A people-centered land transport system must provide for the diverse needs of our society and contribute to a quality, liveable environment especially access to transportation for the low-income groups, the elderly, wheelchair users, families with young children, pedestrians and cyclists (Land Transport Authority, 2008).

17 Ensure Physical Accessibility for All

To make the transport system user-friendly and accessible for all, all new and existing MRT stations will be accessible to persons with disabilities/ impairments. The public bus fleet will progressively be replaced with low-floor wheelchair-accessible buses. Moreover, an island-wide programme was launched to ensure that pedestrian walkways, access to RTS stations, bus and taxi shelters, and all public roads are barrier-free by 2010. Some measures to be implemented include removing obstacles or widening walkways to provide a clear passageway for wheelchair users and using higher reflectivity materials for traffic signs to improve visibility.

18 Affordable Public Transport for Lower-Income Singaporeans

To ensure that public transport remains affordable to low income families, the Government continues to provide targeted help to the needy through Government assistance such as the Workfare Income Supplement Scheme (WIS) and community help schemes such as transport vouchers.

19 Facilitate Cycling

With its increasing popularity, cycling can be a non-motorized transport option to bring commuters to major transport nodes. To facilitate cycling, LTA provide better bicycle parking facilities around MRT stations and bus interchanges; allow foldable bicycles onto buses and trains on a trial basis; close short gaps between the park connectors and transport nodes to cater to commuters who cycle to the MRT stations or bus interchanges; and install appropriate road signs to alert motorists to the presence of cyclists along frequently used routes.

20 Conclusions

In this case, we have described the large scale systems engineering framework and how the conceptualization and systems engineering of Singapore's land transport system were carried out. With the human dimension in centre stage, a people-centered land transport system is being built for all Singaporeans to enjoy an active lifestyle in a vibrant global city. However, this could not be carried out without using a systems approach, understanding the complexities of land transportation and the boundaries faced by the system and integrating them to encourage public transportation.

References

Channel News Asia, Drop in public transport share in Singapore: transport survey (2009), `http://www.channelnewsasia.com/stories/singaporelocalnews/view/1013808/1/.html` (accessed on October 27, 2009)

Haskins, C. (ed.): INCOSE Systems Engineering Handbook: A Guide for System Life Cycle Processes and Activities. International Council of Systems Engineering (2006)

Land Transport Authority, White Paper on "A World Class Land Transport System. SNP Security Printing Pte Ltd. (1996)

Land Transport Authority. Land Transport Masterplan (2008), `http://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/ReportNewsletter/LTMP-Report.pdf` (accessed April 18, 2014)

Land Transport Authority, Household Interview Travel Survey 2012: Public Transport Mode Share Rises to 63% (2013), `http://app.lta.gov.sg/apps/news/page.aspx?c=2&id=1b6b1e1e-f727-43bb-8688-f589056ad1c4` (accessed April 18, 2014)

Land Transport Authority. Land Transport Masterplan (2013), `http://www.lta.gov.sg/content/dam/ltaweb/corp/PublicationsResearch/files/ReportNewsletter/LTMP2013Report pdf` (accessed April 18, 2014)

Land Transport Authority. What We Do (2014), `http://www.lta.gov.sg/content/ltaweb/en/about-lta/what-we-do.html` (accessed April 18, 2014)

Lui, P.C., Tan, T.S.: Building Integrated Large-Scale Urban Infrastructures: Singapore's Experience. Journal of Urban Technology 8(1), 49–68 (2001)

Rittel, H.W.J., Webber, M.M.: Dilemmas in a General Theory of Planning. Policy Sciences 4, 155–169 (1973)

National Climate Change Secretariat, Reducing Emissions in the Transport Sector (2013), `https://app.nccs.gov.sg/data/resources/docs/Consultation%20Paper%20-%20Transport.pdf` (accessed April 18, 2014)

Senge, P.M.: The Fifth Discipline: The Art and Practice of the Learning Organization. Doubleday/Currency, New York (2006)

Sapuan, M.: Managing Singapore's Land Needs. Ethos (2) (2007)

Womack, J.P., Jones, D.T., Roos, D.: The Machine that Changed the World: The Story of Lean Production. Harper Perennial (1991)

Domain-Spanning Change Propagation in Changing Technical Systems

Wolfgang Bauer, Chucholowski Nepomuk, Udo Lindemann, and Maik Maurer

Abstract. Technical products are facing different changes during their lifecycle. These changes can have drastic impacts to the product, either towards its structure or the effort for the implementation. If platform elements used in various products within a product family, the change impact can be capital. Based on existing methods, we developed an approach to propagate changes, triggered by a specific cause, in technical systems. The approach consists of matrix-based methods and allows the propagation via different domains, for example components, functions and variation attributes. Change causes can be linked more easily to the product architecture and the change impact can be traced via different relevant product domains. The approach is illustrated by an industrial case study shown in this contribution.

1 Introduction

Many external causes for product changes exist: country- or region-specific customer needs, legislation, customer habits, competitive products (Pahl et al., 2007) or different market strategies (Porter, 2008). Company internal triggers for changes can be different production sites with different production technologies, raising dynamics of innovations and technology as well as shortened development and release cycles (Ponn and Lindemann, 2011) as reaction to changes the environment.

Theses causes can be triggers for product changes. Often, changes of technical or functional elements can propagate through a product and affect components that are not directly connected to the initiating changed component (Keller et al., 2005). This leads to undetected changes on further component, causing late changes and high costs. To trace and detect undesired change propagations, product architecture offers a good basis.

Wolfgang Bauer · Chucholowski Nepomuk · Udo Lindemann · Maik Maurer
Institute of Product Development, Technische Universität München

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_9

The objective of this contribution is the development of an approach for a systematic analysis of the impacts of technical changes to product architecture. Starting with the cause of a change, its impact to the product architecture is analyzed. The impacts are traced via the domain of functions, components and between these domains. This allows an appropriate mapping of changes to the architecture. Therefore, matrix—based methods from Structural Complexity Management are applied (Eppinger and Browning, 2012; Lindemann et al., 2009).

The paper is structured as follows: starting with an overview of existing methods for the analysis of change propagation in section 2, the developed approach is presented in section 3 and evaluated by an industrial case study in section 4. The paper concludes with a discussion and an outlook for future research.

2 Theoretical Background

Changes in technical systems can be analyzed by using different established methods. A short overview over the most relevant methods for the analysis of change impacts and propagation is presented in this section.

The Change Mode and Effect Analysis (CMEA), presented by (Keese et al., 2006; Palani Rajan et al., 2003), serves the assessment of the flexibility of products. The approach developed by (Köhler et al., 2008), CPM/PDD (Characteristics-Property Modeling/Property-Driven Development), is an approach to illustrate and compare the impact of changes. Another method, based on the CPM/PDD-approach and FMEA, is the Change Impact and Risk Analysis (CIRA), developed by (Conrad et al., 2007). Possible solutions for changes during the product lifecycle are analyzed and assessed regarding risk and impact.

The Change Propagation Analysis (CPA) according to (Giffin et al., 2009) serves the analysis of change requests in order to visualize the changes, the affected components and their role in change propagation.

The Change Propagation Method, developed by (Clarkson et al., 2001), determines the change impact on a product by the means of risk assessment. Based on the CPM, (Keller et al., 2005) developed tools for the visualization of the changes and their propagation.

Koh et al. (Koh et al., 2012) developed a change modelling method (CMM) which is based on the CPM and on the House of Quality (HoQ). The CMM assesses the effects of change options on product attributes by tracing change options via requirements to components.

All of the presented methods assess the impact on the components domain. Some of them use also use other domains for a better traceability of the change within the system. However, the use of indirect change propagation via the chain of change - from the change cause via functions to the components – in an integrated model is missing.

3 Approach for Domain-Spanning Change Propagation

The initial situation underlying this research is an existing product family which should be analyzed regarding the question which functions and components have to be changed due to trends or another expected changes in the future. It should be possible to regard the propagation of any kind of changes, e.g. changes due to failures, changing requirements, new market situations, etc. The change analysis should be conducted domain-spanning, meaning the change path should be traced via functions and components. The resulting change paths should also be visualized to create transparency about the change impact to the technical system.

The applied process is divided into four major steps (see Fig. 1): first step, possible causes for changes are acquired applying the context model (Langer and Lindemann, 2009). These changes are linked to the product architecture using a Multiple-Domain Matrix (DMM) in the next step. In the third step, the effects of the change to the product architecture are propagated across different domains in order to analyze the change impact using change propagation methods in the last step. The four steps are described in detail in the following sub-sections.

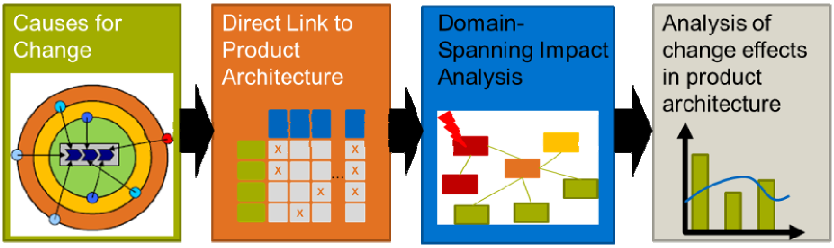


Fig. 1 Approach for domain-spanning change propagation

3.1 Step 1: Causes for Change

In the first step, possible changes are identified. These changes are triggered by influence factors which can emerge in the environment of a company (e.g. legislation, market, competitors) or within the company (organization, processes, resources). The occurring influencing factors are identified by applying the context model of (Langer and Lindemann, 2009). The context model serves the systematic search and documentation of influencing factors and is structured as a search matrix. The context is categorized into following five elements: environment, market, company interfaces, company, and development process. For every context category, classes of influence “technology / knowledge”, “socio-economics”, “politics / legislation”, “resources” are considered. As many influencing factors as possible are gathered by systematically going through the different search fields of the model. It is best practice to involve experts from different fields of the company to achieve a wide range of perspectives and consequently a high number of influencing factors that may trigger changes to the system under consideration.

3.2 Step 2: Direct Link to Product Architecture

The influence factors triggering changes are linked to the product architecture. According to (Ulrich, 1995) the product architecture is defined as "(1) the arrangement of functional elements; (2) the mapping of functional elements to physical elements; (3) the specification of the interfaces among interacting physical components." In this case, the product architecture model is enlarged by the domain "variation attributes". These attributes represent the differentiating parameters of variants derived from the platform. Examples for variation attributes are size, performance, or color.

To create a link between the changes and the product architecture, again Domain-Mapping Matrices (DMM's) (Lindemann et al., 2009) are applied (see Fig 2). In the first DMM, one axis represents the causes, while the other axis shows the components: a change is connected to a component if the occurrence of these causes changes the component in question. The second DMM connects the causes with functions affected by the cause. In the third DMM, the changes are related to the variation attributes if a differentiation attribute of the product is affected in the means of a change by a certain influencing factor. The three DMMs (see Fig. 2) contain the information about the dependency of the causes and their object of change, the product architecture.

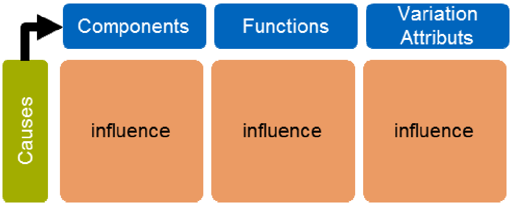


Fig. 2 Matrices for creating the link between causes and the product architecture domains

3.3 Step 3: Domain-Spanning Impact Analysis

In this step, the change impact is propagated to the component domain as components are changed by the engineers. The interfaces have to be designed in a way so that changes do not spread within the whole system. But the creation of the direct link from the changes cause to affected components is not always obvious. It is more likely to identify which of the product's functions or variation attributes are affected by a cause. Therefore, these two domains serve as a translation to the components' domain.

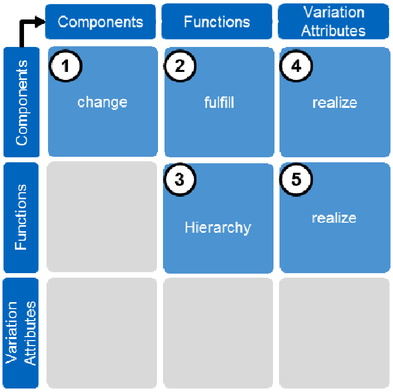


Fig. 3 Product architecture model as MDM-representation

By using the product architecture model (see Fig. 3), the tracing of changes is possible via several domains. The model is represented in a Multiple Domain Matrix (Lindemann et al., 2009) and consists of three domains: components, functions, and variation parameters. The interactions of the components are represented by a bidirectional Design Structure Matrix (DSM) (Steward, 1981) in matrix 1. The type of relations between the physical components is "changes", meaning component A and B are connected if a change of one of these components changes the other component. This relation type delineates geometric contact, signal flow, material flow or energy flow (Pimmler and Eppinger, 1994). The second matrix, a Domain-Mapping Matrix (DMM) represents the link between components and functions. The type of relation is "fulfill". The hierarchic dependencies of the functions are captured in DSM 3 which includes main functions as well as their auxiliary functions. DMM 4 and 5 describe the dependencies between components and functions to variation attributes. Here, a dependency exists if a component/function realizes a variation attribute, e.g. component A realizes a certain performance factor.

The effects of a change can now be traced to the affected components by the DMM's in Fig.2 and navigation through the MDM to the components' domain. The change and propagation mechanism is shown in Fig. 4.

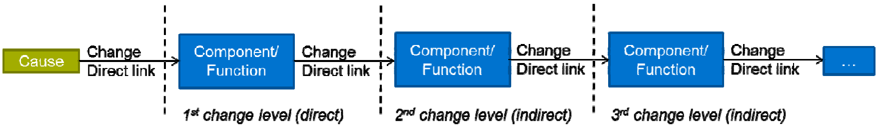


Fig. 4 Change and propagation mechanism

The direct link of change causes to influenced elements (independent form the affected domain) is called "1st level influence and is directly influenced by the cause. All affected elements in the further levels are called according to the number of level, e.g. 2nd or 3rd level elements. These elements are indirectly influenced by the change via elements in higher levels.

3.4 Step 4: Analysis of Change Effects in Product Architecture

After the affected elements on all relevant levels of propagation are known – especially on the component domain – different analyses can be executed. A change propagates either in different domains or in the component domain. Based on this, it can be derived how strong the change impact to the product architecture is. A qualitative estimation of the change effort can be deduced. Another possibility is to calculate structural criteria such as the active and passive sum (Lindemann et al., 2009) of the elements in order to characterize their contribution to the change effect. The active sum of an element indicates how many other elements are influenced by this element. As the active sum gives notice of the change effect of an element to the system, changes on high active elements should be avoided. In contrary, the passive sum of an element describes how many other elements influence this element. Passive elements should be implemented in a robust way, so changes do not spread via these elements as they are sensitive to changes. This approach is similar to (Giffin et al., 2009) where elements are grouped into absorber, carriers or multipliers.

Moreover, the criticality of an element can be calculated by the multiplication of the active and passive sum. The criticality represents the role of an element within the structure based on the in- and outgoing dependencies. The higher the criticality, the more sensitive this element is against changes and can cause numerous changes to other elements, too.

Besides key figures, different visualization methods, e.g. from (Keller et al., 2005) can be applied for the analysis and interpretation of change impacts. Impacts can be visualized by networks or propagation trees. Such networks represent the dependencies between the elements. The elements can origin form different domains, edges represent the dependencies (inter or intra domain). The length of the edges can characterize different information, e.g. the combined risk or the shortest path (Keller et al., 2005).

For change propagation, not only components but also other domains can be used in different change levels. How many change levels are required for the analysis is dependent of the considered system. A domain-spanning propagation tree is used in this approach to visualize the impacts on different domains and different levels. With this representation, it can be determined how many elements are affected in which change level and when the propagation within the system ends.

After analyzing the change impact, the results are interpreted. An important step here comprehends the focus on the elements that are not only possibly affected (via dependencies) but in reality triggered by the change cause. If a change cause is a very general one, numerous elements can be affected in theory. By using design parameters (see e.g. (Ponn and Lindemann, 2011),(Eckert et al., 2004; Ehrlenspiel, 2009)) the user can narrow down the real change impacts. If a pressure or temperature is affected by a change cause, only components (or functions) with corresponding design parameters have to be considered. For conducting such a focusing on the affected elements, search strategies such as feed-forward analysis, impact check list, mine seeking or trace-back analysis (Lindemann et al., 2009) can be applied.

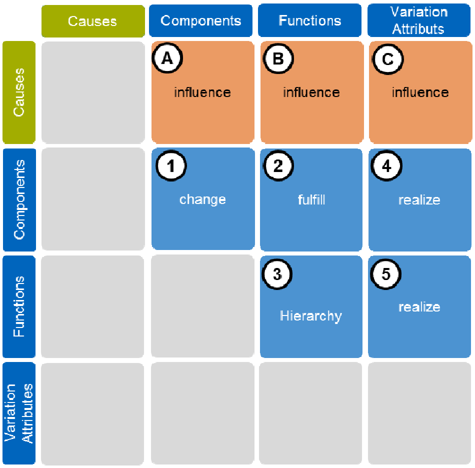
When knowing, which domains and their elements are affected by a change, strategies for handling these changes can be addressed. For example, different design strategies such as size ranges, functional integration or differentiation (Pahl et al., 2007), standardization, platform architectures (Ehrlenspiel, 2009), adaption mechanisms (Kissel, 2012) or the encapsulating of changes into modules by change-driven modularization (Bauer et al., 2013b) can be followed. If the change cause is underlying a cyclical behavior, approaches for handling repetitive tasks can be applied, such as a standardization of the change process for a more efficient execution.

If material, engineering, manufacturing and assembly costs are known, the financial effort of change can be assessed. The cost-effectiveness of a change or and potentials for a re-design can be identified.

4 Case Study for the Domain-Spanning Change Analysis

4.1 Background of the Case Study

The case study is applied on a product family from the whites goods industry. The meta-model in a MDM-representation is given in Fig. 5.



	Causes	Components	Functions	Variation Attributs
Causes		Ⓐ influence	Ⓑ influence	Ⓒ influence
Components		① change	② fulfill	④ realize
Functions			③ Hierarchy	⑤ realize
Variation Attributes				

Fig. 5 Meta-model of the case study

The different domains contain the following number of elements: in total, 127 causes build the domain “causes”. The domain “components” consists of 94 elements. The domain “function” comprises 11 main functions and 46 auxiliary functions, and the domain “variation attributes” includes 4 elements. The relation types between the domains are given in Fig. 5.

The analysis of the use case was classified according to three criteria:

- Component connectivity: a cause is linked to a component with a low, medium and high connectivity
- Number of directly affected elements: low, middle, high number of directly influenced (1st level) elements (functions, components or variation attributes)
- Number of directly affected common elements: high number of functions and components affected by one or more causes.

In the following, the single steps of the presented approach are applied using the case study. Step 1 and 2 are the same for the use cases shown above; step 3 and 4 are described for each of the three different use cases. Due to reason of non-disclosure, the industrial example and its content is presented in an abstract wording.

4.2 Step 1: Causes for Change

In the first step, causes for change for the considered system were gathered. Therefore, the context model (Langer and Lindemann, 2009) was applied in 7 moderated workshops with the industry partner. Each of these workshops, two or more employees from different division (strategic planning, marketing, engineering design, manufacturing etc.) participated. In total, 127 causes were acquired, 98 of them with a direct or indirect link to the technical system (the remaining are e.g. strategic, procedural or organizational causes).

4.3 Step 2: Direct Link to Product Architecture

This step comprehends the direct link of the acquired causes for changes to the product architecture. A detailed process of its acquisition via expert workshops is shown in (Bauer et al., 2013a). All 98 system-related causes were linked to one of the domains “components”, “functions”, or “variation attribute“ using DMM’s. It became true that a direct link to components is not always possible, e.g. as the cause is very abstract. In these cases, especially the access via the product architecture model was very valuable. For example, the cause “development of connectivity of devices” cannot be mapped to components in a direct way, but is related to the function “provide operation”.

4.4 Step 3: Domain-Spanning Impact Analysis and Step 4: Analysis of Change Effects in Product Architecture

In this step, two exemplary cases are presented: on the one hand, the direct link of causes to components; on the other hand, the propagation via the link of causes to functions.

Direct link of causes to components

In this first use case, it is differentiated if a cause directly influences a component with a high, medium or low connectivity. For the assortment of suitable components, the components were ranked according their direct number of component interfaces (see Figure 6). The chosen components are highlighted in the figure.

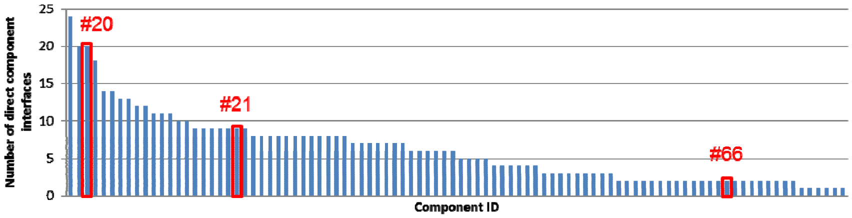


Fig. 6 Component connectivity

Component 20 is an insulation foam material. In the considered case, it is changed by the cause "change of energy efficiency by local regulations". The component has 20 direct interfaces and a criticality of 400, representing a highly connected component. Figure 7 shows the propagation tree for component #20.

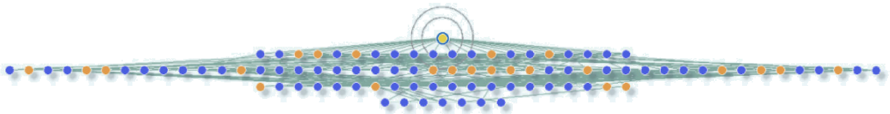


Fig. 7 Propagation tree of component #20

The yellow node indicates component #20, the orange ones represent the affected components, blue the connected but not affected ones. It can be seen that this cause can lead to a change to the whole system already on the 4th level. If component #20 has to be changed because of the cause, already 71% on the second and 92% of the total amount of components on the third level are affected.

For the medium component connectivity, component #21, is chosen. This component has 9 direct interfaces and a criticality of 81. The criticality is over the average criticality of 59. This component can be directly changed by the cause "development of equipment". If component #21 has to be changed, about 70% on the third and 97% of the total amount of components on the fourth level are affected.

Component #66 represents a sensor and a low connected component (two direct component interfaces, criticality = 4). The change triggering cause is e.g. "new development of sensors". Because of the low activity of the considered component, less than one-third of the components are affected on the 2nd level, and the whole system is affected on the 5th change level.

Figure 8 shows the comparison of the change impact of the three considered components with different degrees of connectivity. The propagation on 1st and 2nd level of the high connected component is double then of the medium and low connected components. At the 3rd level, the impacts converge and on the 5th level at latest, the change has affected the whole system in all cases. A change of highly connected components should be avoided or done purposefully to reduce the change propagation effort. If possible, interfaces of such components should be standardized as the change will stop here and the component can be transformed from a multiplier/carrier to an absorber. The direct and indirect change impacts of low connected components are less than medium or high connected ones. Before implementing such a change, it must be ensured to avoid the impacts in the first and second change level to avoid an enormous spreading of the impact. If a change affects a highly connected component in an indirect way on 2nd level or higher, the same procedure can be executed. The result shows that changes on highly connected components lead to more change impacts via propagation. Therefore, components with a high criticality should not be changed or if it is unavoidable changed with a good coordination and a sophisticated impact analysis. Deduced from this result, it is obvious that structural criteria such as the degree of connectivity or criticality are a suitable indicator for change propagation.

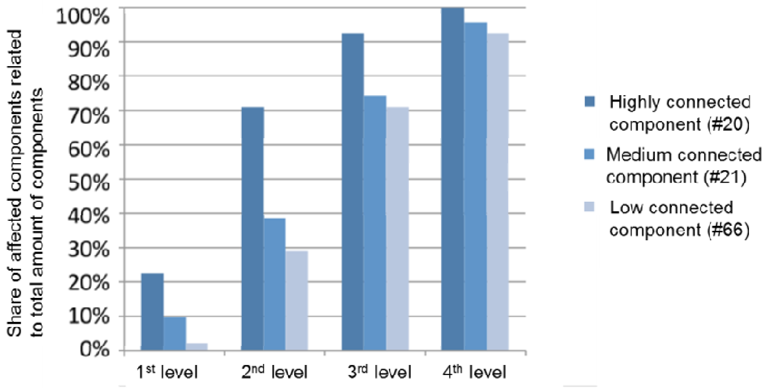


Fig. 8 Comparison of three different connected components regarding their change impact

Indirect link of causes to components

This paragraph gives examples for a direct link of a cause to functions and then propagation to components as the target domain. This is done by tracing the direct linked components to the affected function(s). An example represents the cause “development in the field of consumer electronics” which influences the function “provide operation”. This function is fulfilled by three components (#85: Human Machine Interface, #88: Local Harness, #89: System Harness).

The procedure for propagation is executed in the same manner as in the examples above. The resulting change impact is again in correlation with the connectivity of the starting element.

A cause which affects more than one function is „technological developments (of the core technology of the device)“. 5 of 11 main functions (including 15 auxiliary functions) and in consequence, 46 of 94 components are affected by this cause.

To handle the high possible impacts on the system, the cause can be described in more detail or alternative technical solutions can be derived. It is then possible to isolate the direct link to the main/auxiliary functions, e.g. from 5 to 2 (depending on the situation). Consequently, the different technical alternatives regarding their change impact and consequently the qualitative change implementation effort can be compared.

5 Discussion

The case study shows that the approach is applicable and supports to investigate change propagation via different domains. The underlying architecture models allow an access and propagation between different domains. If a change affects a component, the same procedure can be applied to check if this change affects the functions realized by the component. The procedure can be used iteratively, until all possibilities of change are identified. Therefore, it delivers a solid basis for a change impact analysis and helps to compare different solutions regarding their change effort.

An important aspect regarding a target-oriented propagation is detailing the change impacts. Therefore, design parameters can be included on the edges of the networks to trace the impact more easily. In the presented case, the edges were not further detailed by design parameters as it is a high effort to do so. The cases proved that it is possible for a user of the approach – familiar with the considered system – to trace the change impacts within the system. The dependencies between the system elements help the user to navigate and use them as a kind of check list. By the inclusion of design parameters, the applicability could be enlarged to users not that familiar with the system.

6 Conclusion

The study showed that causes affecting a low number of components, the change propagation through the system is highly depending of the connectivity of the affected component(s). In case of many affected components, the focus of the change and the exclusion of non-changed elements are important to narrow down the analysis.

If a cause is linked to functions, the propagation is dependent of the number of realizing components. The functions can easily be classified regarding their change propagation by calculating their connectivity to components. The procedure for the analysis of domain-spanning propagation is iteratively applicable, independent of the considered domain and uses the connectivity of the elements within the system as a meaningful criterion. Moreover, the results showed that the number of directly influenced elements has an impact on the change propagation, regarding their scope and effort.

The causes were linked to different domains, which allow a reasonable access to the product architecture. Especially the domains “functions” and “variation attributes” are very helpful for creating this link, as a direct link to the affected components is hard to establish.

The reduction of the components from the possible to the actual affected ones was done by a concretization of the change cause, with knowledge of the propagation by the aid of the structural models. Because of the binary information of the structural models, the element connectivity and the change traceability is known.

7 Outlook

The presented approach delivers a qualitative assessment of the change and the consequent effort. Different solution for the implementation of the change can be compared, for example by the number of affected components. To enhance the assessment, the underlying approach and its models can be supplemented by further information: the effort in terms of time and internal costs can be acquired per component or a sum of components (affected by one cause). Therefore, it must be known which department within the company has which amount of effort for implementing the change.

Moreover, the information of the change impacts per cause is currently used to design “change-driven” component modules. Therefore, components are clustered if they are affected by the same cause(s). Mirroring the changes and their temporal behaviour against the company’s future strategy, change modules can be established in the architecture for an efficient change, e.g. for a new product variant, and a fast market response.

Acknowledgments. We thank the German Research Foundation (Deutsche Forschungsgemeinschaft – DFG) for funding this project as part of the collaborative research center ‘Sonderforschungsbereich 768 – Managing cycles in innovation processes – Integrated development of product-service-systems based on technical products’.

References

Bauer, W., Elezi, F., Maurer, M.: Application of DSMs for Analyzing Product and Organizational Structures. In: 15th International Dependency and Structure Modelling Conference, DSM 2013. Hanser Verlag, Melbourne (2013a)

Bauer, W., Elezi, F., Maurer, M.: An Approach for Cycle-Robust Platform Design. In: Lindemann, U., Venkataraman, S., Kim, Y.S., Lee, S.W. (eds.) 19th International Conference on Engineering Design (ICED 2013), Seoul, Korea (2013b)

Clarkson, P.J., Simons, C., Eckert, C.: Predicting change propagation in complex design. In: Design Engineering Technical Conference ASME, Pittsburgh (2001)

Conrad, J., Deubel, T., Köhler, C., Wanke, S., Weber, C.: Change Impact and Risk Analysis (CIRA) - combining the CMP/PDD theory and FMEA methodology for an improved engineering change management. In: International Conference on Engineering Design, Paris (2007)

Eckert, C., Clarkson, P.J., Zanker, W.: Change and customisation in complex engineering domains. Research in Engineering Design 15, 1–21 (2004)

Ehrlenspiel, K.: Integrierte Produktentwicklung, 4th edn. Hanser, München (2009)

Eppinger, S.D., Browning, T.R.: Design structure matrix methods and applications. MIT Press (2012)

Giffin, M., Keller, R., Eckert, C., de Weck, O., Bounova, G., Clarkson, P.J.: Change propagation analysis in complex technical systems. Journal of Mechanical Design 131, 081001 (2009)

Keese, D.A., Takawale, N.P., Seepersad, C.C., Wool, C.L.: An enhanced Change Modes and Effect Analysis (CMEA) tool for measuring product flexibility with applications to consumer products. In: IDETC/CIE. ASME, Philadelphia (2006)

Keller, R., Eger, T., Eckert, C., Clarkson, P.: Visualising change propagation. In: 15th International Conference on Engineering Design, Melbourne, pp. 62–63 (2005)

Kissel, M., Schrieverhoff, P., Lindemann, U.: Design for Adaptability – Identifying Potential for Improvement on an Architecture Basis NordDesign 2012, Aalborg, Denmark (2012)

Koh, E.C., Caldwell, N.H., Clarkson, P.J.: A method to assess the effects of engineering change propagation. Research in Engineering Design 23, 329–351 (2012)

Köhler, C., Conrad, J., Wanke, S., Weber, C.: A matrix representation of the CPM/PDD approach as a means for change impact analysis. In: International Design Conference Dubrovnik, pp. 167–174 (2008)

Langer, S., Lindemann, U.: Managing cycles in development processes-analysis and classification of external context factors. In: 17th International Conference on Engineering Design. Design Society, Stanford University, CA (2009)

Lindemann, U., Maurer, M., Braun, T.: Structural Complexity Management - An Approach for the Field of Product Design. Springer, Berlin (2009)

Pahl, G., Wallace, K., Blessing, L.: Engineering design: a systematic approach. Springer, London (2007)

Palani Rajan, P.K., Van Wie, M., Campbell, M., Otto, K., Wood, K.: Design for Flexibility - Measures and Guidelines. In: International Conference on Engineering Design, Stockholm (2003)

Pimmler, T.U., Eppinger, S.D.: Integration analysis of product decompositions. In: ASME Design Theory and Methodology Conference. Alfred P. Sloan School of Management, Massachusetts Institute of Technology, Minneapolis (1994)

Ponn, J., Lindemann, U.: Konzeptentwicklung und Gestaltung technischer Produkte. Springer, Berlin (2011)

Porter, M.E.: Competitive advantage: Creating and sustaining superior performance. Simon and Schuster (2008)

Steward, D.V.: Systems analysis and management: structure, strategy and design. PBI, New York (1981)

Ulrich, K.: The role of product architecture in the manufacturing firm. Research Policy 24, 419–440 (1995)

Policy Design, Planning, and Management in Global Systems Science

Jeffrey Johnson

Abstract. *Policy Design* is defined to be a new area of inquiry that takes the methods of design into the world of social, economic and environmental policy. Policy exists at many levels and it is increasingly recognized that policies applied to one system may impact on policies applied to other systems. The European Commission suggest a ‘science of global systems’ to improve the way that science can help inform policy and societal responses to global challenges such as climate change, global financial crises, global pandemics, city growth and migration patterns. The new science requires radically novel ideas and thinking to embed scientific evidence into the policy and societal processes. It is here argued that Policy Design in the context of planning and management is an essential part of the methodology of Global System Science.

1 Motivation

Policy Design is defined to be a new area of inquiry that takes the methods and traditions of design into the world of social, economic and environmental policy (Johnson & Cook, 2013). In the sense of Herbert Simon (1969), policy involves a vision of the future as it ought to be, and policy is a science of the artificial. Inescapably, policy is designing the future. Of the many specialist design domains, city planning best exemplifies this. Cities are planned, designed and managed but never finished. They are regulated but not controlled, and their precise state at any point in time is not predictable in the sense of conventional science. The essential feature of policy design is that emerging needs and requirement are satisfied by an iterative process in which possible solutions are generated and evaluated until a satisfactory solution is found or the requirements are reformulated. Design is a coevolutionary process that delivers what we think we want from a process that investigates what is possible. Design is heuristic and cannot guarantee optimum solutions, or even good solutions, but it is the only way we know for creating well-working systems that don’t already exist.

Jeffrey Johnson
Centre for Complexity and Design, The Open University, MK7 6AA, UK

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_10

Policy exists at many levels and it is increasingly being recognized that policies applied to one system may impact on policies applied to other systems. For example, in the UK policies to manage the health and welfare of an aging population have become coupled to policies for managing accident and emergency admissions to hospital. In the face of the economic crisis, cuts to one system have caused a crisis in the other. In another case a European policy to reduce the consumption of fossil fuels by incentivising the production of biofuels in one country caused starvation in another (Foley, 2011).

"Challenges such as climate change, financial crises, or containment of pandemics all suffer from the intrinsic difficulty that they generate strong interdependencies between different social, technological, and environmental systems. When trying to deal with them, different groups tend to address individual systems, rather than multiple interrelated systems, and thereby they typically fail to achieve systemic change. The vision of a science of global systems is that scientific knowledge could act as a catalyst to stimulate creative policy responses to such global challenges, and indeed changes in society in general. A global systems science (GSS) emphasises a 'systems' approach to develop scientific evidence in support of system-wide policy options across different domains. …

GSS also recognises the immense potential for the engagement of civic society throughout the process of decision making by gathering and analysing evidence. A better understanding of this science of global systems will lead to better evidence-based policy decision making." (Bishop et al, 2013).

Design in the context of planning and management is of central importance to the development of Global System Science. Without embracing design the new science of global systems will fail to deliver the holistic solutions so urgently required.

2 Design and the Science of Multilevel Social Systems

Of necessity, Global System Science must be a science of multilevel systems, capable of integrating theories and knowledge from the individual person to the Anthropocene, "the current epoch in which humans and societies have become a global geophysical force" (Steffen et al, 2007). Through many intermediate levels of systems, subsystems and supersystems, almost every individual person has an impact on global systems, and global systems have an impact on almost every individual person.

Policy strives to make global systems behave differently to how they might behave if left alone. Thus the systems created by policy are artificial in the sense of Simon (1969), and they are designed.

The process of designing can be considered to include the generation of new scientific knowledge. For example, there was no science of aeronautics before the invention of flying machines, no science of computer viruses before the invention of computer networks, and no science of human-robot interaction before the invention of robots. We cannot have a science of systems whose components have not been invented.

In this respect the science of social systems is different to traditional science. Whereas the latter is cumulative with new knowledge adding to and reinforcing existing knowledge, human systems become different system when new things are invented. Thus a science of human behavior that was a relevant and 'correct' story of human behavior in the past may not be a relevant and 'correct' story of human behavior today. For example, to what extent is the behavior of people using FaceBook explained by behaviours from the past. Also, some contemporary behaviours were impossible in the past, e.g. crowd-sourced measures of 'liking', and leisure travelling by the masses. Unlike the traditional science, we need custom-designed social sciences to fit the social world as it evolves.

Design is the process that takes new parts and assembles them into new wholes, takes existing parts and assembles them into new wholes, or identifies non-existing parts that are required to complete the construction of new wholes. Design builds the knowledge about the parts and the wholes they form.

Design is the first step in creating scientific knowledge about the relationships between the parts and wholes of artificial systems.

In this respect design is fundamental to a scientific understanding of the multilevel dynamics of the social systems we create and try to manage.

In 2009 an Expert Report of the European Commission asserted that "we have no scientific formalism for representing the bottom-up and top-down dynamics of multilevel systems from micro-levels to macro-levels through meso-levels. This scientific deficit manifests itself across the sciences. In biology there is no formalism able to integrate the dynamics of cells with the dynamics of organs or the dynamics of the whole body. Instead we have many partial models that fit together, at best, descriptively.

In geography and environmental planning we have no formalism that can integrate the choices and behaviour of individuals at the microlevel with the emergence of cities across the globe. In social and political science we have no formalism that can explain why the values and beliefs of individuals aggregate into mutually destructive policies at national level … For all these systems, complex systems science gives reasons why their behaviours are hard to predict. Conventional science assumes that subsystems can be isolated, but complex systems science shows that they may coupled by weak links. This makes subsystems with ill-defined boundaries that are hard to identify and model. These subsystems evolve and coevolve in ways that can only be predicted by modelling their interactions. These interactions do not just occur at particular levels of representation, bottom-up dynamics can cause macroscopic changes, and top-down macroscopic dynamics can cause microscopic changes. ... Creating a formalism for multilevel systems of systems of systems and demonstrating its applicability is on the critical path for science. It is necessary if not sufficient to make progress in many domains. It requires an essential paradigm shift for complex systems science and ICT." (Johnson et al, 2009)

Design has been characterized as building abstract representations of multilevel systems, from the parts to whole (Johnson, 2013, 2014).

At the lowest level are tangible components such bricks and window frames in the design of buildings, or components such chips and capacitors in the design of electrical circuits. In these cases the designer knows the available components at the microlevel and knows the required behavior of the whole system at the macrolevel. Typically a designer conceives subsystems at intermediate levels. These subsystems are abstract and only exist in terms of representations such as annotated drawings and other documentation, or models inside computers. Initially they may be sketches where the parts are not precisely instantiated with existing components, and some components may themselves be sketches of things that do not yet exist. As the subsystems becomes better defined the designer hypothesizes their behavior in terms of their lower level components and the context of higher level assemblies. These hypotheses depend on theories of the system dynamics, and often they are tested by real or computer simulated prototype models. In this respect they resemble scientific experiments in the process of building and testing theories.

In design it sometimes happens that assumptions made about higher level abstract subsystems are incorrect. Sometimes subsystems simply don't fit together because their geometries are incompatible. Sometimes they don't fit together because their interacting dynamics cause problems such as unexpected and unacceptable vibrations. Sometimes they don't fit together because unexpected emergent properties violate the requirements, for example in architecture an important space may be compromised by the noise of air conditioning fans. In such cases the 'theory' of the object being designed did not predict these observations, and as Popper suggested for traditional science, the theory has to be rejected. Since the theory of the design does not match observation, the theory has to be amended. In design this can include imaginative 'fixes' so that local changes in the details of the subsystems can overcome the problem, e.g. a stronger bracket, a higher wall, more memory, or filing down proud edges.

Sometimes these local changes cause unexpected problems elsewhere in the system that may or more not be fixed by local changes. At worst the design has to be abandoned with the loss of all previous work. When this happens late in the design process it can be very expensive in terms of lost work, and it can have severe impact on the schedule leading to expensive delays or even cancellation of the project.

When the design process does not experience such problems, or overcomes them, the design process is one of conceiving intermediate level systems and making them well defined in terms of the components, how those components are to be assembled, the emergent behavior of the intermediate level system, and the dynamics of the interaction of the subsystem with subsystems at low levels, with subsystems at its own level, and with higher level super-subsystems. As the design process proceeds these subsystems become better instantiated. The lower level subsystems are instantiated with real components so that they too become tangible. As these tangible subsystems are assembled higher level intermediate structures become tangible, until eventually the description of the whole system is instantiated with real parts and specifications for their assembly, and it can be built or fabricated.

The construction of the first instantiated system is an experiment testing the hypothesis that it will meet specification. As noted above, if the experiment fails the theory of the design has to be modified. If the experiment succeeds the system has a 'design life' in which evolutionary changes are made to the design, and the life of each individual instance has to be managed from cradle to grave. For example, the owner of a motor car expects the design to include periodic maintenance and procedures to manage the expected or unexpected failure of components, and some countries have regulations for disposal of the end of a product's life. Thus the design of a particular car is part of the design of supersystem that includes sourcing materials and components, manufacturing, marketing, support, and disposal. Such supersystems place topdown constraints on the design of the car, and they too have their dynamics.

3 Design, Planning and Management

In policy, design occurs in the context of strategic or operation planning. Policy Design includes the design of the plan: it establishes how the future ought to be, devises a plan to achieve that future, and manages the implementation of the plan. The plan involves the creation of new systems and subsystems, and these must be designed. After a system or subsystem has been created it needs to be managed, and the design includes determining how this will be done.

Design is the process that begins with knowledge of existing things at the microlevel, defined requirements of a system at the macrolevel, and builds a multilevel representation of a 'possible' system by hypothesising and instantiating subsystems at intermediate meso levels. In the simplest cases this leads to the fabrication of a prototype that can be tested. In manufacturing industries the prototype may the first of millions, e.g. motor cars, it may the first of thousands, e.g. aeroplanes, while in architecture the prototype may be the only one ever built, e.g. iconic buildings such as the Sydney Opera House and the Guggenheim Museum in Bilbao.

In all these cases the construction of the prototype is part of the design process. The experiments on prototypes such as cars and planes test many emergent properties and in the most benign cases lead to small changes that overcome problems or lead to improvements.

Generally design is part of a planning process that begins with an idea of how the system ought to be and initiates the design process to investigate how the system might be. The planning system may even commission more than one design. Once a decision has been made on which design to implement, a plan must be made to implement the design. For products this involves planning production lines, distribution and maintenance systems, and so on. For buildings it involves obtaining consents, preparing the site, commissioning architects to oversee the project, etc., and planning the project in time over months or years.

In cities the planning process provides a context for design. Most cities have planning departments working with other departments to consider the physical infrastructure in the context of how people will function within that infrastructure. The details of the administrative structure differ, but planning departments may have day-to-day management responsibilities for services such as the maintenance of the transport infrastructure and provision of public transportation.

In cities with elected mayors, the incumbent usually strives to make the city as it ought to be, according to their political outlook and constituency. This includes managing day to day things such as public transport at an acceptable level of service but also strategic things such as enabling sufficient economic activity to provide the jobs and the wealth necessary to support the city's tax base. To be re-elected the mayor must deliver sufficiently on the expectations of the voters within a budget that the citizens will tolerate.

Planning and management in cities are characterised by making small local changes in the infrastructure to maintain normal activities, making continuous small or large changes in the context of an existing plan, and occasionally proposing large discontinuous strategic changes as a new ‘masterplan’ for part or all of the city. For example, a run down port may identified for redevelopment, a new transportation plan may address chronic congestion, or a new housing project may address unsatisfied demand for housing. All of these changes at all levels involve design.

At the microlevel urban design can be poor and not serve communities well. In response to this organisations such as the Glass-House (http://www.theglasshouse.org.uk) take a bottom-up approach “supporting and promoting public participation and leadership in the design of the built environment. We provide independent advice, training and hands-on support to community groups and organisations, housing associations, developers, local authorities and other stakeholders, to help them work more effectively together to create better quality places and spaces.”

The mesolevels of city design are typically planned, designed managed by elected councils and mayors, as discussed above. It is being increasingly realized that wealth and wellbeing at national and international level are generated at the level of cities, and that cities perform best when they are autonomous and freed from top-down central control:

“Following a commission from the Prime Minister, Lord Heseltine presented his report No Stone Unturned to the Chancellor of the Exchequer and Secretary of State for Business, Innovation and Skills on 31 October 2012. [It] makes a series of recommendations in all aspects of government policy that affect economic growth. The Government welcomes this report. … The core proposition … is a decentralised approach that breaks Whitehall’s monopoly on resources and decision making, and empowers Local Enterprise Partnerships to drive forward growth in their local areas. … The Government confirms that it is accepting most of the 89 recommendations to dramatically advance the process of decentralisation, unleash the potential of local economies, strengthen partnerships with industry and foster economic growth.” (Her Majesty’s Government, 2013).

Whereas governments can create national environments disposed to the successful development of cities, cities have to compete at a global level. The major cities of the world compete to attract multinational companies from all sectors, and they compete to attract the most talented people worldwide. Thus cities operate in the context of a global systems of cities, looking up to see the macrolevel opportunities and constraints and looking down to manage the meso and micro level needs and aspirations of their citizens.

4 Global System Science

In response to the global financial crisis and the need to manage migration, employment, trade, pandemics, crime, and many other global systems, the European Commission (2013) has identified the need for a Global System Science on which to base policy and action for a wide range of problems affecting European citizens in the short and long terms.

In this context the Commission sought research proposals to “successfully embed scientific evidence in the policy processes for tackling global challenges:

“• Research grounded in theoretical foundations of, among others, systemic risk, decision making under uncertainty or conflicting evidence, mathematics and computer science for Big Data (including their characteristics), algorithmic game theory, cascading/escalating effects in networks, integration and visualisation of Big Data…

Contributions to solving real world problems in one selected problem area - for instance tackling systemic risk in finance/economics, managing growth of cities and migration, or global pandemics – and in particular to tackle cross-cutting policy dependencies and interactions affecting the area of choice.

Novel ideas and technologies to generate and better communicate the scientific evidence-base: advanced simulation of highly interconnected systems; mathematical and tools for analysing (often unstructured) Big Data; integration of the whole spectrum of structure and unstructured data; methods to deal with conflicting data and modeling results; novel data visualisation tools.

Society/human-centred technologies, for instance, new approaches to allow citizens to actively participate in the policy process, to collectively gather and integrate data, analyse evidence, and novel methods to better judge and use scientific evidence: methods, e.g. games, gamification, and narratives to clearly and consistently convey data and modeling results and thereby to stimulate societal responses." (European Commission, 2013).

The expected impact of the programme is research rooted in policy needs, that promotes system thinking, and delivers consistent messages from conflicting data and model results. The research is expected to create a high level of uptake and use of GSS tools and methods in policy and societal processes, including in EC policies. It is also expected to increase the capacity of GSS to help integrate societal responses across policy domains and cross-cutting authorities by development of a system-wide integrated evidence base of data and models.

5 Policy Design, Planning, and Management in Global Systems Science

By definition global systems are multilevel systems and to be realized as well-functioning entities they need to be designed, implemented, and managed.

Global Systems Science cannot just try understand subsystems isolated at local at levels since it must try to understand how these subsystems interact bottom-up and top down from local to global levels. A formalism for representing multilevel dynamics is necessary if not sufficient for Global Systems Science.

The design process is exactly one of creating a formalism to represent multilevel systems, where the dynamics at every level are explicit with known interactions between higher and lower levels. The blueprint for a design cannot have levels missing and interactions ignored.

Global systems such as cities are clearly designed at the micro and meso levels. There is no theory for the emergent behavior of global systems of cities and creating this science is one of the challenges set by the European Commission. If there is some advantageous way of designing the interactions of European cities it is not known. Nor is it known if policies that are advantageous to some cities may be disadvantageous to other cities or other subsystems of 'Europe'.

A major challenge for Global Systems Science is to understand the dynamics of global systems at the macro level and integrate this knowledge with the dynamics of the system at meso and macro levels. Any attempt to test this knowledge by applications involves creating real systems, i.e. designing and implementing them.

A consequence of this is that Global Systems Science is inextricably entangled with design, and that design should be seen as part of the methodology of Global System Science.

6 Conclusions

A new science of Global Systems has been proposed by the European Commission to enable it develop policies in response to the many European problems that involve systems operating at local, national and global levels. This science has to be transdisciplinary and integrative, able to combine heterogeneous knowledge from many source. Of necessity this science will be able to integrate the dynamics of multilevel systems at all levels.

Policy Design is the formulation of visions of hypothetical futures in the context of planning and management. It is argued that the only way to test the new science, and indeed the rationale for it, is the practical applications in which new systems are created based on it. The creation of the artificial is design. Design always involves explicit representation of systems at all levels, where the interactions between micro, meso and macrolevels are well defined and explicit. The design process provides the method to build multilevel systems in a disciplined and replicable way. Thus Policy Design must be recognised as an essential part of the methodology of Global System Science.

References

Bishop, S., Baudains, P., Greenlaw, J., Foden, G., Hunt, J., Johnson, J.: Global System Dynamics and Policy: Best Practice Guidelines Towards a Science of Global Systems (2013), `http://www.gsdp.eu/fileadmin/images/about/global-systems-science/GSDP_Best_Practice_Guidelines_Towards_a_Science_of_Global_Systems.pdf`

European Commission, FET proactive call: 'Global System Science' (2013), `http://ec.europa.eu/research/participants/portal/desktop/en/opportunities/h2020/topics/2074-fetproact-1-2014.html`

H.M. Government, Government's response to the Heseltine Review, Cm 8587 (2013), `https://www.gov.uk/government/uploads/system/uploads/attachment_data/file/221902/PU1465_Govt_response_to_Heseltine_review.pdf`

Johnson, J.H., Bourgine, P., Hales, D.: Complex Systems Science: Expert Consultation Report, Future and Emerging Technolies Proactive, European Commission (2009), `ftp://ftp.cordis.europa.eu/pub/fp7/ict/docs/fet-proactive/shapefetip-wp2011-12-06_en.pdf`

Johnson, J.H., Cook, M.: Policy Design. In: Aiguier, M., et al. (eds.) Complex Systems Design and Management 2013. Springer International Publishing, Heidelberg (2014)

Simon, H.: The sciences of the artificial. MIT Press, Cambridge (1969)

Steffen, W., Crutzen, P.J., McNeill, J.R.: The Anthropocene: Are Humans Now Overwhelming the Great Forces of Nature? Ambio. 36(8), 614–621 (2007), Sciences Module

Foley, S.: Biofuel policy is causing starvation, says Nestlé boss, The Independent (March 23, 2011), `http://www.independent.co.uk/news/business/news/biofuel-policy-is-causing-starvation-saysnestl-boss-2250075.html`

Thinking about the Future of Complex Technological Systems: Which Technologies Should Shape Their Designs?

Jeffrey L. Funk

Abstract. The long lifetime of technological systems increases the importance of understanding those technologies that are experiencing rapid improvements. Ideally we would like to design our systems around these technologies so that we can benefit from the future benefits that these rapid improvements in component performance and cost can provide at a system level. This paper provides data on technologies that are experiencing rapid improvements and it uses transportation systems to demonstrate the impact of rates of improvement on system design issues.

1 Introduction

Many systems are expected to last decades if not centuries. This includes transportation, water and electricity distribution, housing, health care, information, communication, defense, and energy distribution and transportation. During this long life, improvements are expected to occur as new technologies emerge or as improvements to existing ones are made. This suggests that we should design systems around those technologies that are experiencing rapid improvements so that we can benefit at a systems level from these rapid improvements in component performance and cost.

But what are the technologies that are experiencing rapid improvements? We argue that this is not a non-trivial question and that few decision makers can actually answer it. For example, wind turbines are considered an important clean energy technology yet they have only experienced cost improvements of about 2% a year since the early 1980s (IPCC, 2013). Similarly, the energy storage density of Li-ion

---

Jeffrey L. Funk
National University of Singapore, Division of Engineering and Technology Management,
9 Engineering Drive 1, Singapore 117576
e-mail: `etmfjl@nus.edu.sg`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_11

batteries, which impacts on the range and efficiency of electrical vehicles, has only risen at about 5% per year since the early 1990s (Tarascon, 2009). Contrast these rates of improvements with those of electronic components that are between 25 and 40% and one wonders why Li-ion batteries and wind turbines are considered important tools in the clean energy tool chest.

This paper provides data on technologies that are experiencing rapid improvements and the relevance of this data is demonstrated through examples of transportation systems. In doing this, we are not arguing that we have identified all of the technologies that are experiencing rapid improvements or all of their implications for transportation and other systems. We are merely arguing that understanding those technologies that are experiencing rapid improvements is important for designing systems and that better data is needed on rates of improvement in order to better design systems.

This paper first summarizes technologies that are experiencing improvements of greater than 10% per year. Second, the impacts of these rapid rates of improvement on transportation systems are then described. Third, the implications of these analysis for system design are summarized.

2 Technologies Experiencing Rapid Improvements

Table 1 summarizes technologies currently experiencing rapid rates of improvement. This data was found in *Science*, *Nature*, *IEEE*, and other science and engineering journals through extensive reading and searches. We characterize these improvements as annual rates of improvement since most of the time series data are straight lines on a logarithmic plot. We define “currently” as time series data that includes data from the last 10 years. Thus, although some of the time series shown in Table 1 include data from 30 years ago, the technology is included in Table 1 because the times series includes data from the last 10 years and the older parts of the time series data are retained for completeness. The relatively constant improvement rates for longer time periods suggest that these rates can be more reliably extrapolated into the future than can ones with shorter time periods.

The technologies are placed into several categories, which are consistent with characterizations of engineering systems (de Weck et al 2011). These categories include the transforming, storing, and transporting of energy, information, materials, and living organisms. Since a variety of performance measures are often relevant for a specific technology, data was collected on multiple dimensions some of which are represented in performance of basic functions per unit cost while others are in performance of functions per mass or per volume.

As expected, the rates of improvement for information transformation, storage and transmission are experiencing the most rapid rates of improvement. This includes well known components such as integrated circuits and magnetic disk drives and systems composed of them such as computers and mobile phone telecommunication systems. However, newer forms of information processing technologies such as MEMS-based sensors, organic transistors, single walled carbon nanotubes for transistors, and quantum computers are also experiencing rapid improvements and this suggests that there is no end in sight for Moore's Law and improvements in information systems.

Furthermore, many of the technologies classified under energy transformation are actually information technologies. For example, LEDs, OLEDs, GaAs lasers liquid crystal displays, and quantum dot displays transform electrical energy into light and thus are classified under energy transformation. However, unlike engines or motors, these technologies do not do work. Instead, they are used to display or transmit (e.g. lasers) information by transforming electrical energy into light. Even DNA sequencers can be defined as processors of information because they transform biological materials into information. Thus, most of the technologies that are experiencing rapid improvements are applicable to information systems.

The major exceptions are superconductors for energy transmission, carbon nanotubes and grapheme for structural applications, and cellulosic ethanol for energy. The small number of exceptions highlights the need to design systems around information more than other technologies because information technologies experience rapid improvements. Emphasizing information technologies in our design of systems will help us benefit at the system level from improvements in information technologies. We now demonstrate these ideas by looking at transportation systems.

3 Implications for Transportation Systems

The technology currently most discussed for transportation systems is electric vehicles that use batteries and electric motors in places of gasoline and internal combustion engines. The problem with electric vehicles is that the rates of improvement for batteries are very slow, only 5% per year and if these rates continue, the energy storage densities for batteries will not reach the 25 times higher levels found in gasoline for at least 60 years. Since energy storage densities have a large impact on the weight and range of the vehicle, it is very unlikely that battery-powered cars with the range of existing cars will appear for at least 60 years. Furthermore, low energy storage density can lead to a vicious cycle of heavier cars requiring more batteries and more batteries leading to heavier cars.

Of course some may argue that hybrid vehicles are sufficient or that these rates of improvement might increase as scientists and engineers create new materials that have higher energy and power storage densities; these are certainly a plausible future. We argue, however, that these are less plausible futures than the ones we describe below. For the former, users will always prefer a conventional vehicle over a hybrid vehicle since it is much cheaper. For the latter, other technologies are either experiencing more rapid rates than are batteries and/or other technologies are much closer to reaching their necessary levels of performance and cost than are batteries. Furthermore, batteries are not a new technology; they have been used in vehicles for more than 100 years so acceleration in the rate of improvement is unlikely.

Instead, we believe that technologies experiencing rapid improvements (See Table 1) or technologies experiencing moderate rates of improvements but are close to their necessary levels will probably have a larger impact on the effectiveness of transportation systems than will electric vehicles. This section summarizes scenarios that are more plausible than one of electric vehicles reaching the range and cost of conventional vehicles. First, improvements in ICs, sensors, computers, and mobile phones are rapidly improving the effectiveness and usability of public transportation. Specially designed computer systems are improving the capacity utilization of public transportation while improvements in mobile phone technologies and GPS are making it easier to find public transportation such as subway stations and buses. The increased use of GPS in buses will enable users to better monitor the timing of buses and to find the closest bus on their phone.

Second, improvements in cameras, MEMS, lasers, and wireless communication are making autonomous vehicles economically feasible. With annual improvements rates of 25% to 40% for many of the sensors, the cost of the controls for autonomous vehicles will probably drop by 90% in the next ten years thus making autonomous vehicles not much different from conventional vehicles. The largest benefits from automated vehicles will probably occur when roads are dedicated to them and thus tightly packed vehicles can travel at high speeds. Since fuel efficiencies drop as vehicle speeds drop, the use of dedicated roads for autonomous vehicles can have a dramatic impact on fuel efficiency and road capacity, two common problems in most urban and suburban settings. It is likely that autonomous vehicles will become economically feasible before the energy storage densities of current batteries are doubled, which will probably take as long at the last doubling occurred (15 years).

Third, most electrical utilities are combining the Internet, whose performance and cost are experiencing rapid improvements, with the electrical grid to create smart grids. One outcome of adding intelligence to our well established electrical grid can be the capability of vehicles to easily find and purchase electricity from a high density of charging stations in urban and suburban parts of developed countries. Since the cost of distributing electricity is much lower than that of gasoline, the cost of the charging stations is probably not as important as licensing large numbers of firms to sell electricity and thus overcoming the network effects associated with the number of charging stations and electric vehicles. Overcoming these network effects would enable electric vehicles to be charged while a vehicle is parked in a parking garage or along a street in the future (Huber, 2011). This would enable the vehicle to have far smaller storage capacities than are ordinarily thought and thus not depend on improvements in energy storage density.

Fourth, the implementation of densely packed systems of rapid charging stations are also facilitated by the improvements in energy transmission performance that are coming from improvements in superconductors. Superconductors are widely used in magnetic resonance imaging and are beginning to be used in transformers, cables, fault current limiters, motors, generators, and energy storage. We can envision these superconducting transmission lines providing extensive charging points throughout urban and suburban areas.

Fifth, gradual improvements in the performance and cost of power electronics are enabling the “electrification” of automobiles, which reduces the weight and thus the necessary battery capacity of vehicles. This replacement of mechanical controls and drive trains with electrical ones has already occurred in aircraft and heavy trucks and is now occurring in automobiles as the cost of power electronics gradually falls. While the rate of improvement is fairly slow (about 4% per year), announcements by automobile manufacturers suggest that the electrification of vehicles will be largely finished within the next five to ten years (CESA, 2013) and this will reduce the need for large storage capacity in batteries.

Sixth, two alternatives to batteries, capacitors and flywheels, experienced faster rates of improvement in energy storage density than did batteries until 2004 (10% for flywheels and 17% for capacitors) (Koh and Magee, 2006) but more recent data is not available and thus they are not shown in Table 1. This suggests that one of them will eventually have higher densities than do batteries for electric vehicles. Although capacitors have experienced faster rates of improvement than have flywheels, flywheels are currently ahead of capacitors and they are widely used in Formula 1 vehicles, partly because they have higher power densities than do batteries. One of the reasons for the rapid improvements in the densities for flywheels is the replacement of steel and glass with carbon fibers. Carbon fibers have higher strength to weight ratios than do steel or glass and thus can rotate faster than can steel or glass-based ones. Rotational velocity is important because the energy storage density of flywheels is a function of rotation velocity squared. Carbon nanotubes (CNT) have even higher strength-to weight ratios than do carbon fibers and thus CNT-based flywheels can potentially have even higher energy storage densities than do carbon-fiber based ones. Some estimates place the strength-to weight ratios of CNTs at ten times higher than those of carbon fiber. This suggests CNT-based flywheels can have an energy storage density that is ten times higher than that of carbon fiber based flywheels and thus batteries (Krack, Secanell and Mertiny, 2011).

4 Discussion

To the extent possible, systems should be designed around technologies that are experiencing rapid improvements. This paper provides information on technologies that are experiencing rapid improvements and it demonstrated how rates of improvement can impact on one types of system, transportation systems. Transportation systems are expected to be used for decades and it is unlikely that transportation systems designed around electric vehicles will become economically feasible in the near future without considering the low rates of improvement that Li-ion batteries are experiencing. Without considering these low rates of improvement, it is doubtful that electric vehicles with the range and acceleration of existing vehicles will emerge for at least 60 years.

Nevertheless, more data on rates of improvement are needed. There are many types of technologies and many types of systems that can be designed around these technologies. More data will improve our understanding of the tradeoffs between various systems and will thus help us design better systems. We look forward to working with other scholars on better databases of rates of improvement for various technologies and on their impact on various systems.

Table 1 Technologies with Recent Rapid Rates of Improvement

Technology Domain	Sub-Technology	Dimensions of measure	Time Period	Improvement Rate Per Year
Energy Transformation	Light Emitting Diodes (LEDs)	Luminosity per Watt, red	1965-2005	16.8%
		Lumens per Dollar, white	2000-2010	40.5%
	Organic LEDs	Luminosity/Watt, green	1987-2005	29%
	GaAs Lasers	Power density	1987-2007	30%
		Cost/Watt	1987-2007	31%
	Liquid Crystal Displays	Square meters per dollar	2001-2011	11.0%
	Quantum Dot Displays	External Efficiency, red	1998-2009	36.0%
	Solar Cells	Peak Watt Per Dollar	1977-2013	13.7%
		Efficiency, Organic	2001-2012	11.4%
		Efficiency, Quantum Dot	2010-2013	42.1%
		Efficiency, Perovskite	2009-2013	46.5%
Energy Transmission	Superconductors	Current-length per dollar	2004-2010	115%
		Current x length - BSSCO	1987-2008	32.5%
		Current x length - YBCO	2002-2011	53.3%
Information Transformation	Microprocessor Integrated Circuits	Number of transistors per chip/die	1971-2011	38%
	Power ICs	Current Density	1993-2012	16.1%
	Camera chips	Pixels per dollar	1983-2013	48.7%
		Light sensitivity	1986-2008	18%
	MEMS for Artificial Eye	Number of Electrodes	2002-2013	45.6%
	MEMS Printing	Drops per second	1985-2009	61%
	Organic Transistors	Mobility	1984-2007	94%

Table 1 (*continued*)

	Sub-Technology	Dimensions of measure	Time Period	Improvement Rate Per Year
Information Trans-formation	Single Walled Carbon Nano-tube Transistors	1/Purity	1999-2011	32.1%
		Density	2006-2011	357%
	Super-conducting Josephson Junctions	1/Clock period	1990-2010	20.3%
		1/Bit energy	1990-2010	19.8%
		Qubit Lifetimes	1999-2012	142%
		Number of bits/Qubit lifetime	2005-2013	137%
	Photonics	Data Capacity per Chip	1983-2011	39.0%
	Computers	Instructions per unit time	1947-2009	36%
		Instructions per kw-hour	1947-2009	52%
	Quantum Computers	Number of Qubits	2002-2012	107%
Information Storage	Magnetic Storage	Recording density of disks	1956-2007	36.5%
		Recording density of tape	1993-2011	32.1%
		Cost per bit of disks	1956-2007	38.5%
	Flash Memory	Storage Capacity	2001-2013	47%
	Resistive RAM	Storage Capacity	2006-2013	272%
	Ferro-electric RAM	Storage Capacity	2001-2009	37%
	Magneto RAM	Storage Capacity	2002-2011	58%
	Phase Change RAM	Storage Capacity	2004-2012	63%

Table 1 (*continued*)

	Sub-Technology	Dimensions of measure	Time Period	Improvement Rate Per Year
Information Transmission	Last Mile Wireline	Bits per second	1982-2010	48.7%
	Wireless, 100 meters	Bits per second	1996-2013	79.1%
	Wireless, 10 meters		1995-2010	58.4%
	Wireless, 1 meter (USB)		1996-2008	77.8%
Materials Transformation	Carbon Nanotubes	1/Minimum Theoretical Energy for Production	1999-2008	86.3%
	Graphene	Cost per square cm	2009-2013	216%
Biological Transformation	DNA	Sequencing per unit cost	2001-2013	146%
		Synthesizing per unit cost	2002-2010	84.3%
	Cellulosic Ethanol	Output per cost	2001-2012	13.9%

RAM: random access memory; MEMS: microelectronic mechanical systems. Sources: Adapted from (Funk and Magee, 2014)

References

CESA (2013), `http://cesa-automotive-electronics.blogspot.sg/2012/09/dual-voltage-power-supply-system-with.html`

de Weck, O., Roos, D., Magee, C.: Engineering Systems. MIT Press, Cambridge (2011); Renewable Energy Sources and Climate Change Mitigation: Special Report of the Intergovernmental Panel on Climate Change, IPCC 2013. Cambridge University Press (2013)

Funk, J., Magee, C.: Exponential Change: What drives it? What does it tell us about the future (2014)

Huber, P.: Broadband Electricity and the Free-Market Path to Electric Cars, ManhattanInstitute for Policy Research (2011)

Krack, M., Secanell, M., Mertiny, P.: Rotor Design for High-Speed Flywheel Energy Storage Systems. In: Carbone, R., Tarascon, J. (eds.) Energy Storage in the Emerging Era of Smart Grids. Batteries for Transportation Now and in the Future, presented at Energy 2050, Stockholm, Sweden, October 19-20 (2011)

Part II
Posters

A Comparison of DSM and Node-Link Diagram in the Context of Understanding a System and Engineering Change Propagation

Dongwoo Kim and Eun Suk Suh

Abstract. Change propagation management, which is concerned with predicting the propagation of engineering change within a product, is one of the popular areas in engineering. Understanding change is becoming increasingly critical in the prognosis of improvement of system or product. Recently, several approaches for describing changes have been proposed. We focus on the comparison of visualization techniques, Design Structure Matrix (DSM) and node-link diagram in the context of understanding a system and engineering change propagation. In this paper, personal factors such as user's theoretical knowledge and practical experience are considered. This paper will examine which method is more effective for users who should handle change regarding task performance. This work extends the comparison scope of readability to understandability. To test understandability, pens and irons are used as simple examples of systems.

Keywords: Design Structure Matrix (DSM), Node-link diagram, Understandability.

Dongwoo Kim · Eun Suk Suh
Seoul National University
e-mail: `{franciskim,essuh}@snu.ac.kr`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_12

Engineering Change Planning with Consideration of Change Effect on the Project

Duseok Jeong, Haruya Kamiyama, Kazuya Oizumi, and Kazuhiro Aoyama

Abstract. This research aims to provide practical guidance of engineering change for project manager and engineer. Most of products in recent are not developed from scratch. Design based on previous projects or other existing product should be modified to meet new requirements. Likewise, engineering change is one of the powerful driving force and unavoidable work in product development. Therefore, it has been widely researched. However, few researches consider whole development project not only product system, though many previous researches provide the insight into how engineering change effects on the product.

This research's distinctive points are: 1) Changeability of change propagation route ; 2) Evaluate the change plan from perspective of project management.

We assumed that engineer could change the direction of propagation. Actually, an engineer decides where to change and how to change based on the knowledge and experience. Hence, various plans could exist in response to same change request. Each plan has different impact on the development project. This means that the impact on the project from engineering change will vary by the engineer's skill. It means that comparative superiority exists between engineering change plans. Skilled engineer comprehends the change impact quickly and makes efficient plan. However, these processes are still unclear and considered as a tacit knowledge unfortunately. Therefore, this research tried to reproduce skilled engineers' tacit knowledge for provide practical guidance of engineering change.

At first, to describe our assumption that propagation route can be changed, we adopted constraint network from previous research. It enables to deduce various change plans automatically. Then, we made evaluation indices that evaluate change plans from perspectives of Quality, Cost and Delivery indices.

Duseok Jeong · Haruya Kamiyama · Kazuya Oizumi · Kazuhiro Aoyama
Department of Systems Innovation, The University of Tokyo,
7-3-1 Hongo Bunkyoku, Tokyo, Japan, 1138656
e-mail: `jeong@m.sys.t.u-tokyo.ac.jp`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_13

And proposed method validated by using prototype system and case study(build-to-order product). As a result, we concluded that the effectiveness of proposed method seems clear, though there are clearly more work to be researched.

Keywords: Engineering Change, Change Propagation, Design Process and Constraint Network.

From City- to Health-Scapes: Multiscale Design for Population Health

Matteo Convertino, Amanda Aspenson, and Rachelle Schoessler Lynn

Abstract. Reconciling the growing proportion of the global population that lives in urban centers with the goal of creating healthy cities for all poses one of the major public health challenges of the 21st century. Genetics has accounted for only 10% of diseases, and the remainder appears to be from the interaction of multiple socio-environmental causes that potentially determine epigenetic changes leading to diseases. Therefore, quantifying the dynamics of socio-environmental factors and the environment-disease linkages is extremely important for understanding, preventing and managing multiple diseases simultaneously considering population and individual biological information of exposed and non-exposed individuals. This is particularly important for the aim of reprogramming health-trajectories of populations via a quantitative health-based design of cities.

Here we show how complex systems models, and specifically, dynamic network factor analysis (DNF) coupled to global sensitivity and uncertainty analyses can map the exposome-genome-diseasome network (i.e., the macrointeractome), determine network factor metrics useful for urban design, and assess probability distribution of comorbidities conditional to exposure in space and time, respectively. These probabilities are useful to make syndemic predictions by for design of socio-technical and ecological systems and intervention strategies in existing cities via scenario modeling of different design alternatives. As a case study, we use the SHIELD study in Minneapolis focused on measuring children's exposures to multiple environmental stressors and related effects on respiratory health and learning outcomes. Results show the very high degree of directional interaction

---

Matteo Convertino
Environmental Health Sciences & Health Informatics, School of Public Health
Institute on the Environment – Institute for Engineering in Medicine
Institute for Advanced Study – University of Minnesota Twin-Cities
`matteoc@umn.edu`

Amanda Aspenson · Rachelle Schoessler Lynn
MSR Design, Minneapolis, MN

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_14

among exposure factors and their spatial heterogeneity coupled to bi-directionally interacting diseases. We find non-linear conditional probabilities of disease co-occurrence and context-dependent dose-response curves that manifest large health disparities in populations. We show that macro socio-environmental features are much more important than biomarkers in predicting disease patterns with particular focus on respiratory diseases and learning outcomes. This emphasizes the fundamental importance of preventive population health versus downstream ersonalized medicine strategies. Urban texture results as the most important factor, thus, such metric should be clearly considered in the design of socio-environmental systems via a minimization of the systemic health risk.

The developed probabilistic models are extremely flexible for the analysis of big data, city healthscape predictions, and optimal management of communicable and non-communicable diseases in complex socio-ecological systems for systems design. The understanding of linkages between structural, architectural, social, and environmental factors at the population scale will allow designers, architects, engineers, and scientists to design communities - from the material to the city scale - in which population health is the central objective of the design process.

Keywords: complex systems, cityscapes, systemic risk, dynamic network factors, design, respiratory diseases, learning outcomes, urban texture, neighborhood.

Intelligent Taxi Dispatch System with Three Preliminary Rounds of Selection

Jingran Li, Ying Cheng, and Fei Tao

Abstract. Recently, an intelligent taxi dispatch application has become quite popular in China, which has gradually replaced the traditional way. However, through investigation, several disadvantages of the application have been summarized and an improved intelligent taxi dispatch system is proposed that can get rid of the drawbacks of the existing system.

First, through an improved information collection process, unique databases can be established for every passenger, which is the first round of selection. Next, passengers are grouped according to the time when they send order and where they are. By this, the optimization aimed only at individual level has been enlarged to the optimization of the overall passengers in one group. Moreover, with wide use of GPS and powerful compute capability, based on passengers' databases, the second round of selection can be efficiently conducted to find the real-time qualified taxies. Finally, at the third round of selection, in order to adapt this matching problem to a Linear Sum Assignment Problem (LSAP), the final candidate cars in the same number of the passengers can be obtained by a series of proposed evaluation rules.

Considering the specific characteristic of the proposed intelligent taxi dispatch system, the existing algorithm of LASP has been improved to be applied in the model for a better optimization. In the end, the original complex matching problem has been simplified to a simple LSAP which promises less calculation time as well as less passengers' waiting time.

Keywords: intelligent taxi dispatch system, three preliminary rounds of selection, Linear Sum Assignment.

Jingran Li
Student at the School of Astronautics, Beihang University, Beijing, China
e-mail: `yy442037816@163.com`

Ying Cheng · Fei Tao
School of Automation Science and Electrical Engineering,
Beihang University, Beijing, China
e-mail: `chengying268@126.com, ftao@buaa.edu.cn`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_15

Verification in a Future Climate Change Treaty – Technology and Policy Issues

Maria Dominguez and William J. Nuttall

Abstract. Currently Earth Observation systems are driven primarily by two separate government missions: defense and environmental science. The first represents the main observation application for the military and intelligence agencies. Civilian space agencies provide support for environment missions as the main application for their Earth Observation Satellites. Thus far the European Union has led global calls for deep greenhouse gas emissions reduction. Europe has done much to support the development of low carbon energy sources. Future Climate Treaty verification may be provided either by a global monitoring body or via separate regional capabilities. Following an online Delphi consultation undertaken in 2010 we suggest that the EU needs to improve its depth and visibility of thought on this important aspect of climate treaty policy. This work seeks to assess and advance European understanding of these issues as they relate to governments and the aerospace sector. Recommendations are made concerning treaty negotiations, technical matters, European policy and the European space industry.

Keywords: Climate Change, Earth Observation, International Treaties, Satellite Technology.

Maria Dominguez
Judge Business School, University of Cambridge, UK (p/g student 2010-12)

William J. Nuttall
Professor of Energy, The Open University, UK
e-mail: `william.nuttall@open.ac.uk`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_16

Potential Correlation between Uncertainty and Certainty in Complex Systems

Xiaolong Qian and Jincang Zhang

Abstract. The overlap and competition between certainty and uncertainty could be reflected in different complex systems. This article reviews the relevance between certainty and uncertainty in various theories and models. Begin from the nature of the uncertainty principle, the wave function is originated from the uncertainty and implicated the probability, then following the discussion on the details of certainty and uncertainty in the system of folding process of protein, the diffusion limited aggregation model in fractal theory, and a portfolio theory used in super metric space in financial market system. Study on the potential similarity in various theories would help us build more accurate and practical models.

Keywords: uncertainty, competition, complex systems.

Xiaolong Qian · Jincang Zhang
Department of Physics, Shanghai University, Shanghai 200444, China
e-mail: `xiaolongqian@shu.edu.cn`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_17

Financial Feasibility of Introducing Electrical Vehicles into Taxi Service in Singapore

Bo Han, Po-Yu Chen, Yunsha Li, Qi Zhou, and Lijun Tung

Abstract. With a rise in global warming, Governments around the world are turning to electric vehicles as a solution to resolve the partially increasing carbon emissions caused by fuel vehicles. According to International-al Energy Agency (IEA), which is a multi-government policy forum formed by governments from Africa, Asia, Europe, North America as well as EVI, with the aim to facilitate and accelerate the adoption of electric vehicles, the sales of electric vehicles worldwide have more than doubled from 45,000 in 2011 to 113,000 in 2012. However, the challenges to the deployment are the high costs of the electric vehicles, the high initial cost of setting up the charging infrastructures and the great efforts needed to promote electric vehicles to the public.

The purpose of this paper is to analyze the finacial feasibility of introducing electric vehicles to replace diesel taxi for Singapore market. The net present values of electric vehicles and diesel vehicles are discussed and compared on a 10-year (20 cycles) period. After detailed analysis on deterministic base case, base case under uncertainty and base case under uncertainty with flexibility, the recommendation is to introduce electric vehicles to Singapore future taxi market.

This paper presents a four-phase project evaluation technique that provides a higher accuracy in economic analysis of projects. The analysis begins with base case model which utilize discount cash flow (DCF) to compute project net present worth (NPV). The second step takes into consideration possible sources of uncertainty and model them using Monte Carlo simulation. Next step involves identification of possible flexible design in attempt to reduce project impact from market uncertainty. The last step compares results from the previous three scenarios and determines the most profitable design. In this paper, the analysis was

---

Bo Han · Po-Yu Chen · Yunsha Li · Qi Zhou · Lijun Tung
Engineers, National University of Singapore,
#05-408, BLK474 Jurong west street 41 Singapore 640474 Singapore
e-mail: `xshtyhb@gmail.com`,
`{A0039202U,A0122130H,A0122113E}@nus.edu.sg`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_18

performed on the topic to implement electric vehicle into taxis service in Singapore from a taxi company perspective. The analysis covers all four phases depicting a difference in project value at different phase of analysis, when different factors are taken into consideration.

Keywords: Taxi, Cab, Electrical Vehicles, Electrical Cars, Singapore.

Flexibility, Scale, and Value Economics: Mutually Exclusive or Synergistic in City Design?

Sarah B. von Helfenstein

Abstract. Villages within cities – *chengzhongcun, favelas, barong-barongs, barrios marginales, bidonvilles, squatter camps and shantytowns* – extralegal zones that are the shores for tides of humanity congesting cities to the bursting point. Contrast this with the sleek architectural splendor of new city design complete with insect-free green spaces and soil-free urban gardens, the engineering sophistication of 95-story buildings in fantastical shapes and filled with all the luxury goods of the world, the quiet, pollution-less hum of electric vehicles and centralized, unseen water and waste systems, punctuated by the intentional idiosyncrasy of tasteful "pop up" shops and temporary spaces.

Which demonstrates more flexibility, scale, and value? Will cross-fertilization of the two produce further dys-synchrony or unleash the power of flexibility, scale, and value economics?

This paper seeks to "discover what something means by looking at it from the opposite side of the bridge." (de Soto 2000). We examine current trends in the philosophy and design of the complex systems of cities and city infrastructure. We, then, compare these with the world of the *chengzhongcun,* the nano, the deeply sub-system or extra-system. Applying the tools of financial economics to four real world case examples, we present preliminary findings indicating that attention to both scale and flexibility in the design of large complex projects materially affects value. These effects can range from multi-million dollar *negative* net present values to increases in efficiency of over 300% (60+ percentage points). Using inductive reasoning, we draw several general principles from our findings for a proposed paradigm shift in new city design. We conclude our inquiry with some implications for broader issues of the design and management of complex systems and with suggestions for further research.

Keywords: chengzhongcun, new city design, complex urban systems, flexibility, real options, value of complex systems.

Sarah B. von Helfenstein
CEO, Value Analytics & Design LLC
e-mail: `svonhelf@valueanalyticsanddesign.com`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_19

Sensitivity Analysis on System Complexity Index: Considering System Architecture and Different Levels of System Granularity

GwangKi Min, Eun Suk Suh, and Katja Hölttä-Otto

Abstract. In system design, understanding system architecture is an important part of complex system management. System architecture can be properly modeled through decomposition of a system into smaller subsystems, through different level of granularity. Until now, few researchers explored the effect of system granularity on complex systems and its necessity for analyzing the relationship between system granularity and system attributes.

There are various system attributes, such as complexity, modularity and so on. In this paper, system complexity was selected as a system attribute, which will be analyzed in context of system granularity. Overall system structure and interrelationship between various system elements can be explained through this attribute. In this paper, the term 'level of granularity' means the configuration of system elements after system is decomposed. We will focus on the effect of architecture changes.

Various types of system architectures (324 idealized matrices) were created using Design Structure Matrix (DSM) to analyze the effect of system granularity on system complexity. Based on literature reviews of various complexity concepts, system complexity index was selected and verified. Modeled architectures were analyzed to assess their system complexity at two levels of grnaularity using system complexity index. As a result, it was observed that modular system is less sensitive to architectural changes than integral system at the first level of granularity. To verify the results of theoretical works, a complex printing system was analyzed. It was observed that the printing system has similar properties of the ideal integral matrices at the first level of granularity.

Keywords: system architecture, system complexity, system granularity.

GwangKi Min · Eun Suk Suh
Department of Industrial Engineering, Seoul National University,
1 Gwanak-ro, Gwanak-gu, Seoul, Korea, 151-744

Katja Hölttä-Otto
Product Development Pillar, Singapore University of Technology and Design,
20 Dover Rd., Singapore 138682
e-mail: `{minkimkk,essuh}@snu.ac.kr`,
`katja_otto@sutd.edu.sg`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_20

The Changing Face of Singapore's Education System

Eng Seng Chia, Daniel Yeow Chong Ng, Jason Gek Chye Png,
Sean Michael Jurgensen, Chung Siong Tng, and Edwin Wei Quan Toh

Abstract. Education is Singapore's most important investment. Devoid of any natural resources, Singapore needs to focus on the only resource she has – people. The journey from third world to first was premised on being pragmatic; emphasizing centralized control and close alignment between the education and labor market. As such, education policies that focused on sciences, mathematics and English were developed. Over time as Singapore developed, the economy transformed into a knowledge-intensive and innovation-driven one. This forms the impetus for the education system to move away from a rigid, traditional model based on rote learning and move to one that encourages critical thinking and creativity. Against this backdrop, demands for greater inclusiveness and plural political participation echoed an increasing desire for a new social compact in education. Using a systems approach and a large scale systems engineering framework (Chia, 2007), this paper examines the dynamics of Singapore's education system against the historical impetuses for survival, efficiency and catering to individual's ability. It highlights the need for social systems to be flexible to adapt to changing environments and needs.

Keywords: education system, systems approach, large scale systems engineering framework.

Eng Seng Chia · Daniel Yeow Chong Ng · Jason Gek Chye Png
Sean Michael Jurgensen · Chung Siong Tng · Edwin Wei Quan Toh
Temasek Defence Systems Institute
e-mail: `aaron_chia@nus.edu.sg`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_21

Innovations in the Design of Smart Grid Systems for Future Cities

Shisheng Huang, Kristin Wood, Chau Yuen, Kevin Otto, Katja Hölttä-Otto,
Yu Gu, Selin Ahipasaoglu, Saif Benjaafar, Wayes Tushar, Chathura Withanage,
Tao Wang, Anisur Rahaman Molla, Rahul Ashok, and Shaodong Ying

Abstract. The Smart Grid (SG) is the next generation power network composed of intelligent nodes that can operate, communicate, and interact autonomously to efficiently deliver electricity to all stakeholders. It features interconnections of power equipment to enable two-way flow of information and electricity so as to integrate unconventional power sources, manage demand, and efficiently balance supply and demand in real-time.

Consequently, the residential consumer space will be transformed by the introduction of these technologies. The Singapore University of Technology and Design (SUTD) has designed and deployed a residential smart grid tested to conduct end user based experiments to better understand and design this complex system. Here we present insights from the research project on information structure, end user experience research and initial technology development. Through these innovative design insights, there could be potential for great improvements of user acceptance and adoption of these technologies into their spaces.

Particular insights include coverage of residential (private) and shared (public) spaces. As with any new large scale, complex system, there exist many challenges that need to be addressed. Related to these challenges, we introduce efforts of how the test bed at the SUTD has experimented with system architecture, with emphasis on flexibility, privacy and scalability. We find that the information structure should be one of a partially distributed network. We have also initiated research on end user experiences, technologies and expectations in this space.

Studies are being conducted to examine the avenues on how people use electricity coupled with technology development on spatial and temporal tracking of electricity consumption.

Keywords: Smart Grid, system architecture, innovations.

Shisheng Huang · Kristin Wood · Chau Yuen · Kevin Otto · Katja Hölttä-Otto · Yu Gu ·
Selin Ahipasaoglu · Saif Benjaafar · Wayes Tushar · Chathura Withanage · Tao Wang ·
Anisur Rahaman Molla · Rahul Ashok · Shaodong Ying
Singapore University of Technology and Design (SUTD)

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_22

A Systems Approach to Designing Complex Systems: Role of Unpredictability

Rashmi Jain

Abstract. For generations designers have learnt the iterative methods of designing for the purposes of delineating the failure modes that are known to be common in the earlier versions. Henry Petroski and other researchers have argued that designers must go beyond merely ensuring success. They must strive to anticipate the ways in which a design might fail.

Good designers and successful designs try to build in their designs, methods to mitigate such anticipated failures. Anticipating failures is not an easy task, more so, when designing complex systems. US Defense Secretary's comments became famous and gained recognition at several forums when he tried to distinguish between "known unknowns" and "unknown unknowns". While addressing anticipated failures we need to deal with those that are intended and those that are unintended. This paper will cover the topic of identifying unpredictability in complex systems designs and managing its unintended consequences. The author will draw upon her work from systems engineering in defining unpredictability and its sources.

Keywords: Systems Design, Complexity, Unpredictability, Emergence, Unintended, Consequences.

Rashmi Jain
Montclair State University
e-mail: `jainra@mail.montclair.edu`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_23

Integrated Mass Rapid Transit and Bus System in Singapore: MRT Ridership and the Provision of Feeder Bus Services

Devansh Jain and Goh Shu Ting

Abstract. With the aim of improving the quality of life of people of Singapore with provision of better transport services, Land and Transport Authority Singapore recently published its Master Plan 2013. The major objectives mentioned in the plan were to make a comprehensive public transport network with better quality Mass Rapid Transit, bus services along with cycling and walking. MRT is the backbone of the transport system in Singapore, and to promote and increase the MRT ridership, good accessibility to access the MRT stations is a necessity.

The aim of this paper is to investigate the relationship between MRT ridership and the provision of feeder bus services in Singapore planning areas and also to understand the hub and spoke model adopted by Singapore for provision of transport services.

The findings of the study will lead to conclusions made from the Regression model developed by the various factors affecting MRT ridership, and hence will benefit to enhance the services provided by the system.

Keywords: Quality of Life, Public Transport, Mass Rapid Transit, Ridership.

Devansh Jain · Goh Shu Ting
School of Design and Environment – National University of Singapore

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_24

Managing the Complex Systems Development Process with a Collaborative and Open, Experience Based Platform

Eliane Fourgeau

Abstract. Whether for connected cars, smart cities, intelligent buildings, or other industry segments where software dominant and distributed systems come into play, the effective development of complex systems implies mastering multiple disciplines and processes: project, people and data governance processes, requirements based engineering, model based design, test, validation and verification methods …and other key disciplines such as System Architecture with trade-off analysis. All disciplines and processes are interrelated; managing complexity calls for tight collaboration between development stakeholders and demand very high data consistency between system compositional elements, at any level.

Today, organizations are facing significant challenges to maintain coherency, consistency and relevance of the product information across domains as well as to capitalize on it, from one complex project to another; failure to do so though is often resulting in costly reworks, product recalls and worse: attractiveness and competitiveness drops, eventually jeopardized growth.

In nowadays highly competitive, fast changing and demanding markets, saving investment time, resources and money for the creation of higher market appeal products becomes vital.

This presentation will outline the values of an open and collaborative System Engineering Platform, that leverages the richness and efficiency of online communities, the effective composition and reuse of multi-disciplinarily assets and the vision of virtual, model based, user experiences, to facilitate smart systems development. It will also provide references and exemplify how customers have improved their engineering efficiency and carry over capabilities, primarily focusing on Function based development techniques and modern System Architecture development methods that support the process of Systems Engineering.

Keywords: System Engineering, System Architecture, Trade-off analysis, Open System Engineering Platform, Collaborative System Engineering, Platform, Model based Design, Requirement based Engineering, Modeling, Simulation, Modelica, FMI.

Eliane Fourgeau
Dassault Systèmes SA, France
e-mail: `ef7@3ds.com`
`http://www.3ds.com`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_25

Integrated and Sustainable Planning for Industrial Estates

Hsiao Ling Tang

Abstract. Initiated in 1999 as a strategic national development as the leading hub for techno-preneural innovative activities in the region 200ha work-live-play-learn mixed use development, housing biomed, infocom and media industries.

Integrated and Sustainable planning and development strategies for mixed-use, seamlessly connected, compact and liveable Industrial Estates will be shared in this presentation, highlighting JTC's ongoing master planning projects at one north and 2 West.

Keywords: Integrated, Sustainable, Planning, Industrial Estates.

Hsiao Ling Tang
JTC Corporation, TC Summit, 8 Jurong Town Hall Road Singapore 609434 Singapore
e-mail: `hsiaoling.tang@jtc.gov.sg`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_26

Designing Smart Cities – ST Electronics Perspective & Experiences

David Tan Cheow Beng

Abstract. Singapore is a modern metropolis with limited resources. Yet the city-state is fast growing with a vibrancy, while tackling competing demands on its amenities, infrastructure and resources. The Singapore Government has articulated a desire to make Singapore a Smart Nation with knowledge communities driven by intelligence, connected integration and sustained innovation. In essence, a Smart Nation that combines well-thought out policies, a home to educated people and an infrastructure driven by innovative technologies to propel the nation forward into the new millennium.

ST Electronics is one of the largest Information Communications Technologies provider in Singapore. Throughout its journey of success, it has contributed to the well-being of Singapore and beyond. Our touch point involvements range from Intelligent Building Management to Taxi Fleet Management, Intelligent Transportation, MRT Rail Electronics System, Hospital Contact Track and Trace to providing solutions for the Next Generation of Airports and Comprehensive Maritime Awareness systems.

Through our conference presentation : "*Designing Smart Cities – ST Electronics' Perspective & Experiences*", we will share " our experiences and success stories particularly in Complex Airport Hub Operations, Next-Gen Comprehensive Maritime Awareness system and involvement in Safe & Smart Cities programmes both in Singapore and the Middle-East. From conceptualisation to systems design and architecting for business and operational continuity, you'll hear highlights on some of ST Electronics' projects anchoring on Smart Data Analytics, Smart Sensors Grid and Smart Glance for Smart Cities.

Asked to share his vision on Smart Cities, ST Electronics' President, Mr Lee Fook Sun said "*Large cities will be a joy to live in provided they are safe, clean and efficient. Our Smart Cities concept will allow city planners to improve operational efficiencies through better management of water, electricity and gas distribution. We are excited by what technology could offer and we look forward to close partnership to enhance our value proposition.*"

David Tan Cheow Beng
Vice President Business Development,
ST Electronics – Info-Software Systems Pte Ltd tancheowbeng
e-mail: `david@stengg.stee.com`

M.-A. Cardin et al. (eds.), *Complex Systems Design & Management Asia*,
DOI: 10.1007/978-3-319-12544-2_27

Author Index